PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A01H 1/04, 5/00, C12N 5/04, 15/29, 15/64, 15/82, C12Q 1/00

A1

(11) International Publication Number: WO 96/32007

(43) International Publication Date: 17 October 1996 (17.10.96)

(21) International Application Number: PCT/US96/05272

(22) International Filing Date: 15 April 1996 (15.04.96)

(30) Priority Data:
113373 13 April 1995 (13.04.95) IL

(71) Applicants (*for all designated States except US*): YEDA RESEARCH AND DEVELOPMENT CO. LTD. [IL/IL]; The Weizmann Institute of Science, P.O. Box 95, 76100 Rehovot (IL). YISSUM RESEARCH AND DEVELOPMENT COMPANY OF THE HEBREW UNIVERSITY OF JERUSALEM [IL/IL]; 46 Jabotinsky Street, P.O. Box 4279, 91042 Jerusalem (IL).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): FLUHR, Robert [US/IL]; 19/3 Ben-Yehuda Street, 76301 Rehovot (IL). ESHED, Yuval [IL/IL]; 34 Hadror Street, 81523 Yavne (IL). ORI, Naomi [IL/IL]; 39 Hadror Street, 81523 Yavne (IL). PARAN, Ilan [IL/IL]; 6 Harakefet Street, 99797 Carmei Yossef (IL). ZAMIR, Daniel [IL/IL]; 10 Derech Eretz, 70700 Gedera (IL).

(74) Agent: BROWDY, Roger, L.; Browdy and Neimark, Suite 300, 419 Seventh Street N.W., Washington, DC 20004 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: TRANSGENIC TOMATO PLANTS CONTAINING A FUSARIUM RESITANCE GENE

(57) Abstract

The invention provides genes from the *I2* Fusarium resistance locus of tomato belonging to a multigene family herein designated I2C. The DNA molecules of the invention are useful as a tomato resistance gene to plant vascular diseases caused by Fusarium pathogens, particularly *Fusarium oxysporum* f.sp. *lycopersici* race 2, or as probes for breeding Fusarium-resistant tomato lines or for screening of new diseases in plants of the Solanaceae family. Further provided are Fusarium-resistant tomato lines transformed by an I2C resistance gene of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Transgenic Tomato Plants Containing a Fusarium Resistance Gene

FIELD OF THE INVENTION

The present invention relates to genes from the *I2* Fusarium resistance locus of tomato belonging to a multigene family herein designated I2C, useful either as a tomato resistance gene to plant vascular diseases caused by Fusarium pathogens, or as probes for breeding Fusarium resistant tomatoes or for screening of new diseases in related plants of the Solanaceae family, and to transformed plants, particularly Fusarium resistant tomatoes.

BACKGROUND OF THE INVENTION

Resistance to pathogens is thought to involve a specific recognition between a resistant plant and the pathogen, which triggers a set of responses that act to confine the pathogen. The specificity of this process is considered to involve a recognition between the products of a plant resistance (R) gene and a cognate pathogen avirulence gene (Dangl, 1995; Staskawicz et al., 1995). The characterization of resistance genes is of major importance for elucidating the initiation of the cascade of events that leads to specific resistance responses, as well as for more efficient introduction of resistance to pathogens into important crops.

Several resistance genes have been cloned recently by positional cloning or by transposon tagging. These genes include: the *HM1* gene of maize (Johal and Briggs, 1992), the *Pto* gene of tomato (Martin et al., 1993), the *Cf-9* gene of tomato (Jones et al., 1994), the *RPS2* (Bent et al., 1994; Mindrinos et al., 1994) and the *RPM1* (Grant, 1995) genes from *Arabidopsis* , the *N* gene from tobacco (Whitham et al., 1994), and the *L6*

gene from flax (Ellis et al., 1995; Lawrence et al, 1995). These resistance genes show diverse biological characteristics. The *HM1* gene is the only example to date where the gene product acts directly to inactivate a component of the pathogen attack, or a compatibility factor (Briggs and Johal, 1994). The other genes belong to a different genetic category, that of incompatibility (or gene for gene) interaction, based on the recognition by the resistance gene product of an avirulence (or incompatibility) component of the pathogen, which does not necessarily participate in the compatibility or in the infection processes (Briggs and Johal, 1994). These genes are all involved in resistance processes characterized by hypersensitive response (HR). In spite of their origin from different plant species, and their divergent specificity to viral, fungal or bacterial pathogens, a group of these R genes share several structural features. A nucleotide-binding domain (P-loop) and five additional amino-acid stretches of unknown function are conserved in their N-terminal region. A region of leucine-rich repeats (LRR) is present in their C terminus, though the consensus sequence and the length of the repeats are different among them. LRR were shown to be involved in protein-protein interactions in other proteins (Kobe and Deisenhofer, 1994; Kobe and Deisenhofer, 1995), and may have similar role in resistance genes. The *N* gene, the *L6* gene and the *Cf*-9 gene were shown to belong to large gene families, partially clustered with the resistance gene. The detailed genomic distribution of these multigene families is yet unknown.

The soil-born fungus *Fusarium oxysporum* is the causative agent of severe wilt diseases in a large variety of plant species world-wide. It is an imperfect fungus for which no sexual cycle is known. The tomato-specific pathogen *Fusarium oxysporum* f. sp. *lycopersici* (*F.o.l*) causes the disease Fusarium wilt. The fungus penetrates the vascular system of roots from both resistant and susceptible varieties, mainly through wounds. During a compatible interaction, which leads to disease, the fungus proceeds through the vascular system which eventually collapses. This leads to wilt and often to death of the plant. During an incompatible interaction, resulting in resistance, the fungus is confined to the lower part of the roots, and further symptoms do not develop. Several mechanisms, not including HR, were suggested to be involved in this resistance. They include: the production of inhibitory secondary metabolites, and structural barriers such as vascular gelation, callose deposition, and abnormal membrane outgrowths of vascular

parenchyma cells, termed tyloses. Most of these processes, thought to be involved in resistance to vascular diseases, are detectabe also in compatible interactions, though to a lesser extent. Therefore the exact sequence of events that leads to resistance is still unknown.

Three races of *F.o.l.* and their cognate R genes have been identified in tomato. The classification of different *F.o.l.* isolates into races does not correlate with their general genetic resemblance, as established by restriction fragment length polymorphism (RFLP) analysis and distribution into vegetative compatibility groups (VCG; Elias et al., 1993). The *I* locus, introgressed from *L. pimpinellifolium*, confers resistance to *F.o.l* race 1, and is located on the short arm of chromosome 11, between the RFLP markers *TG523* and *CP58* (Eshed and Ori, unpublished). The *I3* locus from chromosome 7 of *L. pennellii* confers resistance to races 1, 2 and 3 of *F.o.l.* (Bournival et al., 1990). This locus appears to be composed of three separate but linked genes (Scott and Jones, 1991). The *I2* locus, introgressed from *L. pimpinellifolium*, confers resistance to race 2 of the pathogen. We previously mapped *I2* to the long arm of chromosome 11, between the RFLP markers *TG105* and *TG36*, very close to *TG105* (Segal et al., 1992; Ori et al., 1994). In previous studies we utilized recombinant inbred (RI) lines for mapping *I2* (Ori et al., 1994). However this population turned to be problematic for mapping of this region because of a very high recombination rate, including double recombinations, especially in the region of *I2*.

SUMMARY OF THE INVENTION

It has now been found in accordance with the present invention that high resolution genetic and physical mapping of the *I2* region, using a large and conclusive F2 population (3200 meiotic gametes), show complete cosegregation between *I2* and a cluster of genes on chromosome 11 belonging to a new multigene family, herein designated I2C.

Additional multigene family members are dispersed between four different loci, on three different chromosomes, either in clusters or as single genes. Two I2C genomic clones were isolated from the locus completely linked to *I2* and sequenced, and were herein designated *I2C-1* and *I2C-2*. Their sequences show striking structural similarity with a group of recently isolated resistance (R) genes, which includes the above-

mentioned *RPS2* and the *RPM1* genes from *Arabidopsis*, the *N* gene from tobacco and the *L6* gene from flax. These genes confer resistances to specific pathogens of viral, bacterial and fungal origin, and share common features. They contain a conserved nucleotide binding domain, termed P-loop, in their N terminus, and five other conserved domains of unknown function. At least half of their C terminus is composed of leucine rich repeats (LRR).

A few partial cDNA clones from the I2C family were further examined, such as the herein designated I2C-3 and I2C-4 cDNA clones, and show that family members differ from each other mainly by insertions or deletions.

The deduced amino acid sequence encoded by members of this gene family reveals a region of LRRs, as well as a P-loop and other motifs in common with the above-mentioned recently characterized plant resistance genes.

Thus, in one aspect, the present invention provides a DNA molecule selected from the group comprising:

(i) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-1 (SEQ. ID. NO.:1);

(ii) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-2 (SEQ. ID. NO.:2);

(iii) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-3 (SEQ. ID. NO.:3);

(iv) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-4 (SEQ. ID. NO.:4);

(v) a DNA molecule characterized by containing a coding sequence representing at least 60% similarity with the encoded open reading frame in the DNA sequence of at least one of the DNA molecules (i) and (ii);

(vi) a DNA molecule capable of hybridization with any one of the DNA molecules (i)- (v) under moderately stringent conditions;

(vii) a DNA molecule that differs, by insertion, deletion or as a result of the degenerative nature of the genetic code, from the DNA sequences (i)-(vi); and

(viii) a fragment of any of the DNA molecules (i)- (vii).

The DNA molecule defined in (v) above contains preferably a coding sequence representing 70-80% similarity with the encoded open reading frame in the DNA

sequence of at least one of the DNA molecules (i) and (ii). The moderately stringent conditions required in (vi) above are such as those conditions described in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd. edition, Cold Spring Harbor Laboratory Press, New York, 1989.

One of the members of the multigene family I2C consisting of a DNA molecule as defined in (i)-(vii) above will confer resistance to *Fusarium oxysporum* f.sp *lycopersici* race 2 in tomato plants. In another aspect, the invention relates to a gene construct comprising such DNA as a genomic clone including regulatory sequences that flank the coding region thereof, and to a cosmid, into which said gene construct has been subcloned, for direct transformation of tomato plants.

In another embodiment, a DNA molecule according to the invention may be subcloned into a plant transformation vector under the control of regulatory elements capable of enabling the expression of said DNA molecule in plant cells. Said DNA regulatory sequences comprise, for example, a plant promoter, a DNA sequence that enhances translation of the mRNA transcribed from said DNA molecule and a polyadenylation/terminator sequence.

In a further embodiment the invention provides a tomato cell line or a tomato plant line transformed with a cosmid or with an expression vector of the invention, and to tomato plants regenerated from said transformed cells

In another aspect of the invention, a DNA molecule or fragment thereof as defined in (i)-(viii) above may be used as a direct RFLP probe employing standard protocols for breeding tomatoes resistant to *Fusarium oxysporum* f.sp *lycopersici* race 2, or to examine the homologous multigene family in related plants of the Solanaceae family, e.g. potato, pepper, petunia, eggplant, preferably plants which have colinear genomic maps with tomatoes, for finding new species-specific disease linkages with said probes. The thus bred tomato plants and related plants of the Solanaceae family are also encompassed by the present invention.

Thus the invention provides a method of selective breeding of Fusarium resistant tomatoes employing a DNA molecule according to the invention as a direct restriction fragment length polymorphism (RFLP) probe, which comprises:

(i) marking said DNA molecule with a suitable marker; and

(ii) reacting said probe of (i) with DNA extract of a tomato plant under hybridization conditions;

thus obtaining restriction-length polymorphism that is indicative of a resistance-type gene, which facilitates the selection of progeny that contains said resistance gene.

The invention further provides a method of screening new diseases in plants of the Solanaceae family employing a DNA molecule according to the invention as a direct RFLP probe which comprises:

(i) marking said DNA molecule with a suitable marker; and

(ii) reacting said probe of (i) with DNA extracts from said Solanaceae species,

thus identifying the homologous gene family in those species which can be linked to known resistance genes in those species.

DESCRIPTION OF THE DRAWINGS

Figs. 1A-C depict genetic and physical maps of the tomato gene *I2* region. (1A) Genetic linkage map of chromosome 11, adopted from Eshed et al. (1995). The *I* and *I2* Fusarium resistance loci were positioned according to Eshed (unpublished) and Segal et al. (1992), respectively; (1B) High resolution mapping of the genetic region spanning RFLP markers *TG105A* and *TG36*, as revealed from analysis of 1600 F2 and F3 individuals; (1C) Physical map of YAC 340-63, with relevant markers indicated. The total length of YAC 340-63 is 350 kb.

Fig. 2 shows Southern blot analysis of genomic tomato DNA of resistant and susceptible parental types and of the fixed recombinant F2 plant BR 30(5). DNA samples were digested with Taql, and the blot hybridized with SL8 probe. R and S indicate *F.o.l* race 2 resistant and susceptible individuals, respectively. R lanes, parental types which are a nonrecombinant resistant type F2 individual from the F2 population initiated from Br5577; S lanes, parental lines which are the sensitive tomato inbreds *L. esculentum* var. M82 and *L. esculentum* var. S-365. BR 30(5) is the single recombinant identified within the *SL8* cluster, from the entire F2 population. TR1-TR8 indicate resistant-type polymorphic bands, as established by examination of this and additional gels. Sizes in kb are indicated on the right.

Figs. 3A-B show distribution of *SL8*-homologues in the tomato genome. (1A) Linkage maps of chromosomes 8, 9 and 11. The linkage maps were adopted from Eshed et al. (1995). The relevant introgressed regions of the ILs are illustrated on the right of each tomato linkage map (solid lines). Asterisks indicate approximate map positions of known disease resistance genes. The mapped positions of the *SL8* loci are indicated; (1B) Southern blot of TaqI digested DNA of representative ILs. *L. pennellii* fragments in the blot are designated A, B, C, D and E according to their genomic location, as indicated in panel A.

Fig. 4 depicts the nucleotide sequence and deduced amino acid sequence of the clone herein designated I2C-1 [SEQ. ID. NOS:1 and 5, respectively]. The first translated nucleotide is no. 1. Sequences conserved between resistance genes are double underlined. Leucine Rich Repeats (LRRs) region and other AA repeats are single underlined. A putative Leucine-Zipper domain is underlined with dots.

Fig. 5 depicts the nucleotide sequence and deduced amino acid sequence of the clone herein designated I2C-2 [SEQ. ID. NOS:2 and 6, respectively]. The first translated nucleotide is no. 1. Sequences conserved between resistance genes are double underlined.

Fig. 6 depicts the partial nucleotide sequence of the 3' of the cDNA clone herein designated I2C-3. [SEQ. ID. NO:3]

Fig. 7 depicts the partial nucleotide sequence of the 3' of the cDNA clone herein designated I2C-4. [SEQ. ID. NO:4]

Fig. 8 shows comparison of the deduced amino acid sequences of the genomic clones I2C-1, I2C-2 [SEQ. ID. NOS:5 and 6] and of the resistance genes RPS2 and RMP1 from *Arabidopsis*, N from tobacco and L6 from flax (Bent et al., 1994; Dangl, 1995; Grant, 1995; Jones et al., 1994; Mindrinos et al., 1994; Lawrence et al., 1995; Whitham et al., 1994) [SEQ. ID. NOS:7, 8 and 9]. Residues numbers are from the first translated methionine of each sequence. Consensus sequence in the N terminal region is indicated only when minimum number of gaps was needed for alignment of at least 5 out of 6 residues. Symbols are: con., consensus sequence; a, aliphatic residue; - and +, negatively or positively charged residues, respectively. Boxes containing stretches of conserved residues are indicated by a line above the sequences, and numbered from I to VI.

Fig. 9 shows alignment of the leucine reach repeats (LRR) of I2C-1 [SEQ. ID. NO:5], and their consensus sequence.(Top) Alignment in the region from residues 558 to 1220 of I2C-1 where alignment to consensus sequence is optimized; (Bottom) A comparison of the consensus sequences of the I2C-1 LRR with those of the resistance genes RPS2, N and Cf-9, and the T-LR SAG expression site associated leucine rich protein from *Trypanosoma brucei*.α represents an aliphatic residue.

Fig. 10 shows comparison of the 3' end of four *I2C* family members. *I2C-1* and *I2C-2* are the deduced amino acid sequences [SEQ. ID. NOS:5 and 6] as in Fig.8. *I2C-3* and I2C-4 are the deduced amino acid sequences [SEQ. ID. NOS:10 and 11] derived from partial cDNA clones from a λgt 10 library. The sequence of *I2C-4* is a chimera between three ORFs, originally separated by one base insertions which caused two frame shifts. The junctions where separated ORFs were combined are indicated by arrows. Con. indicates consensus and is shown when a residue is present in all 4 sequences. Numbers are from the first methionine in sequences *I2C-1* and *I2C-2* [SEQ. ID. NOS:5, 6] and from the first residue of the available sequence for *I2C-3* and *I2C-4* [SEQ. ID. NOS:10 and 11]. Brackets indicate a repeat unit of 23 amino acids which appears in variable copy number in the 3' end of the cDNA clones.

Fig. 11 depicts cosmids 12-134 and 12-150, which contain the genes I2C-1 and I2C-2, respectively, in the BamHI site of cosmid TDNA 04541.

Fig.12 shows sense constructs from the I2C-134 cosmid, which contains the genomic clone I2C-1 prepared in the PGA492 binary vector.

Fig. 13 shows sense (2-1 and 5-1) and antisense (6-3 and 31-17) constructs comprising the genomic clone I2C-1 or the cDNA clone I2C-3 for transformation of Fusarium resistant plants.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to a multigene family, I2C, which is dispersed on three different chromosomes of the tomato genome. Two lines of evidence suggest that a member of this gene family is the *I2 Fusarium* resistance gene. The first is the complete cosegregation of some of the genes from this family with the *I2* gene; the second is the striking structural similarity between members of this family and a group of recently isolated plant resistance genes (Bent et al., 1994; Dangl, 1995; Grant et al., 1995; Jones

et al., 1994; Mindrinos et al., 1994; Lawrence et al., 1995; Whitham et al., 1994). In addition, *I2C* genes from the *SL8D* locus of the family, which maps to *I2*, were shown here to be very highly polymorphic between *F.o.l.* resistant and susceptible varieties. In a similar fashion, the *Pto* resistance gene was also shown to be a gene family highly polymorphic between bacterial speck resistant and sensitive varieties (Martin et al., 1993). Which of the family members are responsible for resistance can be determined by extensive complementation tests with all members of the *SL8D* cluster.

The I2C gene family contains a few motifs that have been identified in plant resistance genes. The N terminus contains a P-loop and 5 additional conserved boxes of unknown function. Different classes of P-loop motifs are common to many but not all nucleotide binding proteins (Saraste et al., 1990) . The consensus of this motif varies significantly between different classes of nucleotide binding proteins, but is highly conserved within each class. The consensus sequence GMGGaGKTT, where a designates an aliphatic amino acid, is highly conserved among the P-loops of the genes *I2C-1*, *I2C-2*, *RPS2*, *RPM1*, *N* and *L6*. No other protein in the gene bank were found to contain this consensus sequence. In contrast to the *Cf-9* gene, and perhaps the *RPS2* gene, the *I2C-1* deduced protein sequence does not predict any membrane-spanning domain. Residues 623 to 645 of *I2C-1*, included in the LRR region, fit the consensus for a leucine zipper (Busch and Sassone-Corsi, 1990) .This motif is considered to be involved in dimerization of DNA binding proteins. However, sequences that fit this consensus are abundant in the databanks, and the existence of this consensus does not necessarily imply a function.

In common with recently isolated plant resistance genes, the C-terminal parts of the *I2C* genes are composed of leucine rich repeats. The LRR consensus comprising 23 amino acids, together with a lack of a membrane spanning domain in the gene, are consistent with an intracellular location of this gene family (Jones et al., 1994). The LRRs of four members of the I2C family show high homology to each other and differ from each other mainly by insertions or deletions. This may be indicative of evolutionary processes, and hint at mechanisms that generate new diversity of LRR. Interestingly, one of the cDNA analyzed, *I2C*-4, contains stop codons within the LRR region, and therefore may result in a truncated protein. This is reminiscent of phenomena described for the *N* and the *L6* genes, where truncated transcripts were described, apparently arising fr m alternative splicing. In the case of *I2C-4* this mechanism is less likely, since the genes

analysed thus far from the family appear to lack introns. The I2C LRR-region consensus is homologous to the trypanosome variable surface glycoprotein (VSG) expression site associated gene *T-LR*. This gene is thought to be involved in the regulation of adenylate cyclase function (Ross et al., 1991; Smiley et al., 1990). LRR were described recently in many proteins, and may be involved in protein-protein interactions (Colicelli et al., 1990; Kobe and Deisenhofer, 1994). Thus, the LRR region may be responsible for specificity of interaction, either with a protein component from the pathogen, or with downstream factors involved in signal transduction. The crystal structure of an interaction of an LRR containing protein, an RNAase inhibitor (RI), with RNAase, was recently described (Kobe and Deisenhofer, 1995). The RI contains a horse-shoe like structure in which individual, 28-29 residues long repeats constitute b-a hairpin units which are aligned parallel to a common axis (Kobe and Deisenhofer, 1993). However, the LRR consensus of *I2C* and of the other plant resistance genes differ from that of RI. In addition, the repeat length of I2C varies between 19 and 32 amino acids per repeat. Similar variation in repeat length can be observed in other R genes, which may imply a less organized or different structure than that found for RI.

Using novel tomato genetic populations (Eshed et al., 1992; Eshed and Zamir, 1994; Eshed and Zamir, 1995), all members of the *I2C* family have been mapped according to the present invention. The *I2C* genes are distributed to five locations in the genome, two of which are clusters of several genes, both on chromosome 11. Some of the recently isolated resistance genes were also shown to be members of large, clustered, gene families, but complete mapping data for them is lacking. The complex pattern of distribution of *I2C* is remindfull of the case of *L* and the *M* rust resistance genes of flax (Ellis et al., 1995). In that case, *L* appears to be a single multiallelic gene, whereas homologous sequences map to a more complex *M* locus containing a gene cluster. Both loci may contain resistance genes specific for different races of the same pathogen, or to different pathogens. It is interesting to note in this respect, that the *I2C* copy from chromosome 9 (*SL8B*) maps with a resolution of 5 cM to both the *Tm-2a* TMV resistance gene (Young et al., 1988), and the *Frl Fusarium oxysporum f.* sp. *radicis lycopersici* (*F.o.r.l.*; Laterrot and Moretti, 1995). The *I2C* cluster *SL8C* maps in the vicinity of the *Sm Stemphylium* resistance gene, with a resolution of 10 cM (Behare et al., 1991). However, no member of the *I2C* family maps to the *I* locus on the short arm of

chromosome 11, or the *I3* locus in chromosome 7 (Bournival et al., 1990; Scott and Jones, 1991), which confers resistance to races 1, 2 and 3 of *F.o.l.*

Considering that a member of the *I2C* family encodes for the *I2* resistance gene, the present invention shows commonalties between a wilt disease resistance gene and other disease resistance genes. Despite the lack of HR in vascular disease resistance, the I2C family belongs to the superclass of resistance genes described for leaf HR. This raises questions concerning the role of the various functional domains of R genes in upstream and downstream events that result in different types of resistance mechanisms.

For transformation of tomato plants, a genomic I2C clone according to the invention may be subcloned into any suitable cosmid, such as cosmid TDNA 04541 (Jones et al., 1992). Such constructs contain more than a few kb of genomic DNA upstream and downstream from the gene coding region sufficient for regulated expression. These constructs can be used for direct DNA transfer into plant cells by electroporation (Dekeyser et al., 1990); by polyethyleneglycol (PEG) precipitation (Hayashimoto et al., 1990), by balistic bombardment (Gordon-Kahn et al., 1990), or by Agrobacterium-mediated transformation (Jones et al., 1992).

Other engineered constructs according to the invention comprise a DNA molecule of the invention and DNA regulatory elements enabling the expression of said DNA molecule in plant cells . Said DNA regulatory sequences comprise, for example, a plant promoter, a DNA sequence that enhances translation of the mRNA transcribed from said DNA molecule and a polyadenylation/terminator sequence.

The plant promoter used in the invention is selected from tissue specific and non-tissue specific plant promoters of different kinds, derived from both mono- and dicotyledoneous plants. The preferred promoter is the commercially available cauliflower mosaic virus (CaMV) 35S promoter that is generally expressed in most, if not all, plant tissues, including vascular tissues. Another example of promoter expressed in vascular tissues that can be used in the invention is PRB-1b (Eyal et al., 1993).

The promoter is to be found in the 5' region of the gene. At the 3' end of the promoter, a short DNA sequence for 5' mRNA non-translated sequence may be added to enhance translation of the mRNA transcribed from the gene, such as the omega sequence derived from the coat protein gene of the tobacco mosaic virus (Gallie et al., 1987).

Downstream at the 3' end of the resistance gene DNA coding sequence a terminator DNA sequence containing the 3' transcription termination and polyadenylation signal of the mRNA from the resistance gene is installed. Terminator DNA sequences comprised within the 3' flanking DNA sequences of any cloned genes can be used, such as the 3' untranslated sequence of the octopine synthase gene of the Ti plasmid of Agrobacterium tumefaciens (Greve et al., 1983), or more preferably the 3' untranslated sequence of the nopaline synthase gene (Depicker et al., 1982).

The gene constructs of the invention can be subcloned into expression vectors, such as the Ti plasmids of Agrobacterium tumefaciens, the preferred plasmid being the pGA492 binary vector (An., 1986).

The expression vector comprising the resistance gene is then introduced into plant cells by a transformation protocol capable of transferring DNA to dicotyledoneous plant cells, preferably by infection of plant cells with Agrobacterium tumefaciens using the leaf-disk protocol (Horsch et al., 1985). For this purpose, tomato leaf explants are infected and the transformed tomato cells are cultured on a suitable medium, preferably a selectable growth medium. Tomato plants can then be regenerated from the resulting callus. Tissue cultures of transformed tomato cells are propagated to regenerate differentiated transformed whole plants. Transgenic plants are thereby obtained whose cells incorporate a Fusarium resistance gene in their genome, said gene being expressible in the cells. Seeds from the regenerated transgenic plants can be collected for future use. Transformed plants that are resistant to Fusarium oxysporum f.sp. lycopersici race 2 can be selected by incorporating a selectable marker such as resistance to kanamycin.

The DNA molecules of the invention can further be used as markers in selective plant breeding as direct RFLP probe as exemplified in Examples 2 and 3 and Figs. 2 and 3 herein, or small fragments thereof can be used for PCR-based technology in marker assisted breeding (review by Tanksley et al., 1995).

In the RFLP probe technology, the DNA from different varieties of tomatoes or related Solanaceae plants is digested by restriction enzymes and fractionated on an agarose gel. Digests are chosen such that the gene of interest will be polymorphic. The DNA fragments are transferred to a nitrocellulose or other similar blotting agent and hybridized to the I2C probe of the invention, preferably under conditions of 6xSSC, 0.5% SDS, at 65°C. The blots are washed, preferably at 2xSSC at room temperature, and

subjected to autoradiography. Further washings at higher temperatures and lower SSC concentrations can be carried out for higher stringency.

The individual DNAs in the population that show the polymorphic signature of the resistance gene can then be further used in a breeding program for the desired traits.

An advantage of using these DNA sequences as direct RFLP probes in selective plant breeding is that the presence of disease resistance in plants can then be examined without using phytopathological methods. In addition, by using a direct or tightly-linked DNA marker as probe (based on RFLP or PCR-based technologies), it is possible to select for the desirable trait, i.e. resistance gene without accompanying genetic drag, i.e. transfer of the desired trait by breeding without incorporation of flanking unwanted traits.

The DNA molecules of the invention can further be used as probes to identify homologous multigene families conferring resistance to different diseases in related plants of the Solanaceae family, such as those which have colinear genomic maps with tomatoes, e.g. potato, eggplant, pepper, petunia and the like, for example, by using a marker of the I2C-1 family as a DNA probe to clone the related gene family from said related Solanaceae species, and then using that clone or I2C members directly as DNA probes to analyze genetic lines of those species for RFLP linkage to resistances in said species.

In the PCR-technology, a preferred approach uses specific oligonucleotides synthesized according to the sequences of selected regions of the I2C family such that the fragment generated in a PCR reaction will yield a polymorphic band for the resistance trait either by being specific for the resistance-type gene and therefore yielding a null PCR reaction in sensitive plants, or by yielding PCR fragments of different size or different restriction patterns upon using sensitive or resistant sources of the DNA.

The invention now will be illustrated according to the following non-limiting Examples and the drawings herein.

EXAMPLES

In the Examples, the following Materials amd Methods are used.

MATERIALS AND METHODS

(a) Plant material and genetic mapping

Four *L. esculentum* segregating F2 populations were used for genetic mapping. In each case, an initial cross was made between a parent resistant to *F.o.l* race 2 (R), and a susceptible parent (S). The resistant and susceptible parents for the first 3 populations were, respectively, c.v. Motelle and c.v. Money Maker; c.v. Mogeor and c.v. Vendor; and c.v. Motelle and LA 1113 (chromosome 11 marker stock, kindly provided by Dr. C.M. Rick (UC Davis). The fourth population was initiated from the commercial hybrid line Br5577 (AB Seeds, Ness Ziona, Israel), as F1. The results obtained from the 4 populations were pooled.

(b) Genomic and cDNA libraries, plasmids and probes.

YAC8 (YAC 340-63, Cornell collection), which contains the RFLP marker *TG105*A, was generated from the *F.o.l* race 2 resistant tomato line Rio Grande - PtoR, and cloned in the vector pYAC 4 (Martin et al., 1992). Probes from YAC 340-63 that were used for the genetic and physical mapping are as follows. D2 is from a genomic lambda library of the tomato line VFNT cherry, selected during chromosome walking from *TG105*. SL8, SR8 and 6-16, are subclones of lEMBL3 clones from a library of the yeast line that contains YAC 340-63. D14 is a cDNA clone selected by YAC 340-63 from a cDNA library from roots of the *F.o.l.* race 2-resistant tomato *L. esculentum* c.v. Mogeor. Additional clones, previously described (Ori et al, 1994), were not polymorphic or not informative in these populations.

cDNA clones which represent members of the I2C family were isolated from three different cDNA libraries. The cDNA libraries were all constructed in λgt10, from roots or leaves of resistant-type *L. esculentum*. Positive clones were equally abundant in the three different libraries. While large (>3 kb sizes) clones have been previously isolated from these libraries, all the SL8 clones were partial and contained only the 3' end of the genes. Cosmid clones were isolated from a genomic library of the *F.o.l.* race 2 resistant variety *L. esculentum* var. Mogeor, constructed in the cosmid TDNA 04541 (Jones et al., 1992).

(c) Sequence analysis

Cosmid clones were either subcloned into the Bluescript plasmid or sequenced directly. The cDNA clones were subcloned into the Bluescript vector and then sequenced. Sequencing was performed with an automated Applied Biosystems Sequencer. Sequence analysis was performed using the sequence analysis software package of the 'University Wisconsin, Genetics Computer Group' (Devereux et al, 1984).

(d) Physical mapping

Yeast DNA for pulse field gel electrophoresis analysis was digested with limiting amounts of the restriction enzymes MluI, XhoI and SalI, to obtain successive partial digestions. The digests were fractionated on counter clamped homogeneous electric field (CHEF) gels (BioRad), blotted and hybridized with probes. The maximal distance between a pair of markers was estimated according to the smallest partial band that contained both markers. Additionally, the DNA was digested with the rare cutters SgrAI and PmeI.

(e) Genetic mapping

RFLP analysis and *F.o.l.* inoculation were performed as previously described (Segal et al., 1992). F2 plants (1200) were screened for recombinants between *TG105* and *TG36*. When necessary, F3 seedlings from the recombinant plants were screened in order to fix the recombination to a homozygous state. An additional 400 F3 plants were screened, and more recombinants were identified in the region. All the recombinants were analyzed for the different RFLP markers and for Fusarium resistance.

Example 1. High resolution mapping of I2

The *I2* Fusarium resistance gene was previously mapped to chromosome 11, between RFLP markers *TG105A* and *TG36*, 0.4 centimorgan from *TG105A* (Figure 1A, and Segal et al., 1992). To obtain higher resolution mapping of *I2*, we generated new markers in the region of *TG105A*, by chromosome walking from *TG105A* on lambda clones, and by subcloning a 350 kb YAC clone, YAC 340-63, that hybridized to *TG105A*. Pulse Field Gel Electrophoresis (PFGE) of YAC 340-63 was used to physically position genetically informative markers, as shown in Figure 1C. In order to localize the position of *I2* relative to the new markers, a segregating population of 1600 plants (F2

and F3) was screened for recombinations between *TG105A* and *TG36*, and 57 recombination events were detected. The recombinant plants were then tested for *F.o.l* race 2 resistance, and for RFLP markers located between *TG105A* and *TG36*. According to the resulting map (Figure 1B), *I2* maps to the multi-copy marker *SL8D*, which represents the edge of YAC 340-63 (Figure 1C) and lies genetically between markers 6-16 and *TG36*, 0.23 cM from 6-16 and 1.3 cM from *TG36*. In previous studies we utilized recombinant inbred (RI) lines for mapping of *I2* (Ori et al, 1994). Interestingly, in spite of some inconsistencies in linearity of markers between the RI and the F2 population described here, *SL8D* completely cosegregated with *I2* also when mapped using the RI lines (data not shown). The *SL8* marker showed a remarkably high rate of polymorphism between *F.o.l* race 2 resistant and susceptible lines. In comparison, flanking markers showed a much lower degree of polymorphism, as judged by the paucity of restriction digests that yielded polymorphic bands.

Example 2. *SL8* is a member of a gene family cosegregating with *I2*

The complete genetic cosegregation of *SL8D* with the *I2* resistance gene and the unique level of polymorphism between resistant and susceptible lines prompted us to further characterize the multicopy marker *SL8*. Sequence analysis revealed that SL8 contains an open reading frame with similarity to a group of recently isolated resistance genes (see below). This suggested that SL8 is a part of a gene that belongs to a family which includes the *I2* resistance gene. The gene of which the SL8 probe was the 3' part was therefore designated *I2C-1* (I2 candidate 1).

We wished to further characterize the different SL8 family members as RFLP markers, draw criteria to distinguish between them, and analyze their genomic distribution. A comparison of the *SL8* RFLP patterns of resistant and susceptible type lines, obtained with the restriction enzyme TaqI, is shown in Figure 2. Approximately 17 different TaqI bands hybridized to the SL8 probe, and many of them were polymorphic between resistant and sensitive lines. Resistant-type bands, consistent among all tested lines, were designated TR1-TR8. The rest of the bands were either nonpolymorphic or polymorphic between the susceptible lines, because of the different parental origins. As several polymorphic TaqI bands were detected with the SL8 probe, direct allelism cannot be established. In the entire F2 population, a single recombinant plant was identified

between all the polymorphic SL8-bands (BR 30(5) in Figure 2), which separated *SL8D* into two distinct loci, *SL8D-1* and *SL8D-2*. (Figure 1B). Except for this case, all polymorphic SL8 bands cosegregated. The recombinant individual BR30 (5) is sensitive to *F.o.1* race 2, and contains one resistant-type TaqI band, TR7, but lacks the others (TR1-6 and TR8). Similar additional southern blots and progeny tests of BR 30(5) confirmed these results (data not shown). Therefore, bands TR1-TR6 and TR8 appear to completely cosegregate with each other and with *I2*, and are all candidates for the resistance gene. However, the possibility of a recombination within the gene should also be considered. This could result in a sensitive plant containing a part of the resistance gene, and consequently a polymorphic band that belongs to the resistance gene. In addition, the possibility exists that a nonpolymorphic band represents the resistance gene. The latter is unlikely as will be shown below.

Example 3. Genomic distribution of I2C

As only a subset of the *SL8* copies showed polymorphism between the parents of the F2 populations used for mapping, additional populations were incorporated to map all the SL8 fragments. The first is an introgression-lines (IL) population, in which single chromosome segments from *L. pennellii* were introgressed into a *L. esculentum* background. Both parental species of the IL population carry the susceptible allele at the *I2* locus (Eshed and Zamir, 1994). Figure 3A illustrates the genomic segments introgressed from chromosomes 8, 9 and 11 in the IL lines, which proved relevant for the *SL8* mapping. All *SL8* copies appear to be polymorphic between *L. esculentum* and *L. pennellii* (Figure 3B, lanes 1 and 2), a feature which facilitated their mapping. DNA digests from the IL lines were compared by southern blot hybridization with SL8 with that *L. esculentum* and *L. pennellii* (Figure 3B). *L. pennellii* bands that are contained in each IL, as well as their allelic *L. esculentum* bands that are absent from these lines, represent SL8 copies that originate from the region introgressed in the respective line. One *SL8* copy mapped to the short arm of chromosome 8 (*SL8A*), as one *L. pennellii*-type band is present, and one *L. esculentum* -type band is absent from IL 8-1. Similarly, one copy with weaker homology mapped near the centromere of chromosome 9 (*SL8B*; Figure 3A,B). More accurate location of these two copies was obtained using lines containing shorter introgressed segments of the region, derived by selection of

recombinants from the F2 of the original IL crossed back to *L. esculentum* var. M82 as illustrated in Figure 3A. The rest of the *SL8* copies mapped to the long arm of chromosome 11. Two of the introgression lines, IL 11-3 and IL 11-4, contain *L. pennellii* segments from the long arm of chromosome 11 (Figure 3A). By comparing these two lines, three genetically distinct groups of SL8 family members could be identified on chromosome 11. The first (*SL8C*) maps to the region exclusively introgressed in IL 11-3, the second (*SL8D*) to the region of overlap between the IL 11-3 and IL 11-4, and the third (*SL8E*) to the region exclusively introgressed in IL 11-4. As previously established, the *I2* resistance gene maps to the region of overlap designated *SL8D*.

Higher resolution mapping of the chromosome 11-based *SL8* loci was accomplished using an F2 population of 150 plants, generated from an initial cross between *L. esculentum* and an introgression line that contains the long arm of chromosome 11 (line 11, Eshed et al., 1992). Analysis of the F2 population corroborated the division of SL8 markers into clusters. *SL8C* and *SL8D* cosegregated completely with the RFLP markers *TG546* and *6-16*, respectively, and *SL8E* mapped between markers *TG26* and *TG105*, 0.25 cM from *TG26* (Figure 3). The susceptible-type *L. esculentum* is a common parent between the IL population and the F2 population. A comparison of the SL8 RFLP patterns of the ILs (Figure 3) with those of resistant and susceptible *L. esculentum* plants (Figure 2), shows clearly that nearly all fragments that belong to group *SL8D* are polymorphic between the resistant and susceptible F2 parents (compare Figure 2 and Figure 3). The nonpolymorphic bands in Figure 2 belong mostly to the other groups. This indicates that the region containing cluster *SL8D* is the region which was originally introgressed from *L. pimpinellifolium* into *L. esculentum*.

Example 4. Heterogeneity in recombination rates in the I2 region

The locus *SL8-D*, containing at least 4 clustered members of the I2C family, spans a 0.03 cM region in the *I2* locus. Two cosmid clones of approximately 20 kb insert from this cluster, I2C-134 and I2C-150 (Fig. 11), contain only one copy of SL8. Hence, 0.03 cM genetic interval spans at least 20 kb, which assigns an estimation of at least 670 kb/cM in the region containing this group. This ratio is similar to the average of 550 kb/cM over the entire genome. In contrast, in proximal regions the ratio is <150 kb/cM between D2 and SL8 (*I2C-1*; Figure 1) and approximately 43 kb/cM between *TG105A*

and *TG26* (Segal et al., 1992). High variability in the physical to genetic ratio is common in the context of different regions of the chromosome, and is shown here to fluctuate regionally as well.

Example 5. *I2C* Genes share structural similarity with a family of plant resistanc genes

Candidate members of the *I2C* gene family were isolated from genomic libraries of *F.o.l* race 2 resistant tomatoes utilizing SL8 as a probe. The isolated clones were compared to the genomic DNA on southern blots, and clones that contain resistant-type polymorphic bands from the *SL8D* locus were further characterized. The cosmid clone I2C-134 (Figs. 11, 12) contains the *I2C-1* gene, that includes in its 3' region the SL8 marker (Fig. 1C). It was found to contain the polymorphic bands TR1 and TR5 (Fig. 2). I2C-134 also exhibits resistant-type polymorphic bands after digestion with other endonucleases, such as HindIII, DraI and EcoRI (data not shown). Cosmid I2C-150 (Fig. 11) contains another gene, *I2C-2*, represented by the polymorphic bands TR4 and TR7 (Fig. 2). As the recombination in individual BR 30(5) has occurred between band TR7 and the other resistant-type bands, it would be expected that the gene which contains TR7 will not contain other resistant-type bands. The presence of both TR4 and TR7 bands in I2C-150 could be explained either by a recombination within the gene *I2C-2* in the individual BR 30(5), or by comigration of bands of different origin. Other cosmids isolated contained non-polymorphic bands, and were not sequenced.

One continuous open reading frame was identified in each of the genes *I2C-1* and *I2C-2*. Fig. 8 shows a comparison between the deduced amino acid sequences of the *I2C-1* and the *I2C-2* genes and recently isolated plant resistance genes. Although the overall homology is rather low, an intriguing structural similarity is apparent. All genes contain in their N terminus a conserved nucleotide binding site, P-loop, and other conserved amino acid stretches of unknown function, which are shown in Fig. 8 as boxes I-VI. In their C terminus they all display a long region of leucine-rich-repeats (LRR), which spans at least half of the gene. The LRR of *I2C-1* are aligned in Fig. 9. The N terminal parts of the *I2C* repeat segment are comparable to the consensus LRR found in the resistance genes *RPS2*, *N* gene and *L6*, and to the consensus of the *T-LR* VSG expression site associated gene from *Trypanosoma* (Fig. 9, Ross et al., 1991; Smiley et al., 1990).

The latter protein shares 52 % similarity and 25 % identity with the 3' part of *I2C-1*. The C-terminal parts of the repeat segments are not conserved, and are of variable length.

Example 6. Transcribed sequences from the I2C gene family contain insertion and frame-shifts.

In order to compare different resistant-type members of the I2C family, three different cDNA libraries of resistant tomato varieties were screened with the SL8 probe. Fifteen independent clones were isolated, and all of them were shorter than *I2C-1*, containing only the 3' ends of the genes. The reason for obtaining only partial clones is not known, as much longer inserts have been isolated from these libraries. Two of the longest clones, designated *I2C-3* and *I2C-4*, of 1200 and 1600 bp long, respectivley, were sequenced. Interestingly, one of the cDNA clones, *I2C-4*, contains two frame shifts, and thus, if translated, would result in a truncated peptide. In Fig. 10, *I2C-4* is artificially shown as a continuous chimera of the 3 different ORFs, and the junctions are indicated with arrows. The comparison of the 3' region of the two genomic and two cDNA clones is shown in Fig. 10. Striking insertions or deletions, chiefly in the C-terminal region, can be observed. Most insertions are shared by at least two different genes, though the combinations are different for each insertion. Interestingly, close to the C-terminus the different genes contain between 2 and 6 repeats of an almost identical sequence. These repeats are indicated in Fig. 10. The largest insertions in genes *I2C-3* and *I2C-4* are made up exclusively of these repeats.

Example 7. Plant transformation procedures

In order to correlate the disease resistance capacity of the I2C genes, they have been transformed into tomato plants in a few different formats:

(a) Cosmid clones 12C-134 and 12C-150 containing the complete inserts of clones I2C-1 and I2C-2, respectively, in the BamHI site of cosmid TDNA 04541 (Jones et al., 1992) (Fig. 11), were directly transformed into sensitive tomato lines VF-36 and Money Maker (Jones et al., 1992). Another 10 cosmids have been isolated by homology to I2C and were similarly transformed. To this end the cosmids were transferred into Agrobacterium tumefaciens using standard transformation procedures. The binary vector

A. tumefaciens LBA 4404 is suitable for the transformation procedure. Tomato explants are inoculated as described in Jones et al., 1992.

(b) Similarly, clone I2C-1 was introduced into the PGA492 binary vector supplemented with the B-domain of the 35S promoter (constructs 134 A and 134H) and were directly transformed into the same sensitive tomato lines as described above.

Fig. 12 depicts sense constructs from the I2C-134 cosmid containing the I2C-1 clone in the PGA492 binary vector. A depicts an AccI subclone, 134A, containing approximately 3kb upstream and 300 bp downstream to the open reading frame (ORF), cloned into the ClaI site of the PGA492 vector; B. depicts a HincII subclone, 134H, containing around 3kb upstream and 800 bp downstream to the translated region, cloned into the HpaI site of the PGA492 vector; D depicts a small subclone from the BS subclone of cosmid 134 depicted in C, containing 300 bp upstream and 800 bp downstream to the translated region, cloned downstream from the B domain (B Dom) of the 35S promoter, in the PGA492 vector, to create 134B.

(c) Antisense and sense partial clones were constructed with partial sequences from I2C-1, I2C-2, and I2C-3, and subcloned into 35S omega expression vectors in the PGA492 binary vector. These constructs (2-1, 6-3, 5-1 and 31-17) were transformed into resistant tomato lines (Motelle) as described in section (a) above.

Fig. 13 depicts the antisense and sense constructs 2-1, 6-3, 5-1 and 31-17 for transformation of Fusarium resistant tomatoes. On the top is a map of SL8-134 (I2); indicated is the HindIII fragment used for the antisense and sense cloning in 2-1 and 6-3, which spans nucleotides 2540-3716 in the I2C-1 sequence shown in Fig. 4. Constructs 5-1 and 31-17 contain the full-lenght insert of cDNA I2C-3. 35S is the cauliflower mosaic virus (CaMV) 35S promoter, Ω is a translation enhancer from tobacco mosaic virus, and nos 3' is a terminator, 3' untranslated sequence of the nopaline synthase gene. All the clones were prepared in the PAG492 binary vector using the unique XbaI site.

The resulting transformed tomato lines are tested for complementation of Fusarium resistance in sensitive lines like Money Maker or abrogation of Fusarium resistance in the Motelle line. Tests are carried out by inoculating 10 days old seedlings with freshly prepared Fusarium fungus cultures and disease is estimated during 10-20 days following inoculation. Sensitive plants show retarded growth browning of vascular tissue and usually dye within 20 days.

REFERENCES

An, G., (1986) *Plant Physiol.* **81**, 86-91.

Behare, J., Laterrot, H., Sarfatti, M., and Zamir, D. (1991). Restriction fragment length polymorphism mapping of the *Stemphylium* resistance gene in tomato. *Mol. Plant-Microbe Interact.* **4**, 489-492.

Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B.J. (1994). RPS2 of *Arabidopsis thaliana:* A leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856-1860.

Bournival, B.L., Vallejos, C.E., and Scott, J.W. (1990). Genetic analysis of resistances to races 1 and 2 of *Fusarium oxysporum* f. sp. *lycopersici* from the wild tomato *Lycopersicon pennellii*. *Theor. Appl. Genet.* **79**, 641-645.

Briggs, S.P., and Johal, G.S. (1994). Genetic patterns of plant host-parasite interactions. *TIBS* **10**, 12-16.

Busch, S.J., and Sassone-Corsi, P. (1990). Dimers, leucine zippers and DNA-binding domains. *TIG* **6**, 36-40.

Colicelli, J., Field, J., Ballester, R., Chester, N., Young, D., and Wigler, M. (1990). Mutational mapping of RAS-responsive domains of the *Saccharomyces cerevisiae* adenylyl cyclase. *Mol. Cell. Biol.* **10**, 2539-2543.

Dangl, J.L. (1995). Piece de Resistance: Novel classes of plant disease resistance genes. *Cell* **80**, 363-366.

Dekeyser, R.A. et al., (1990) *Plant Cell* **2**, 591-602.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: Transcript Mapping and DNA Sequence. *J. Molec. Appl. Genetics* 561-573.

Devereux, J., Haeberli, P. and Smithies O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395 .

Elias, K.S., Zamir, D., Lichtman-Pleban, T., and Katan, T. (1993). Population structure of *Fusarium oxysporum* f. sp. *lycopersici*: restriction fragment length polymorphisms provide genetic evidence that vegetative compatibility group is an indicator of evolutionary origin. *Mol. Plant-Microbe Interact.* **6**, 565-572.

Ellis, J.G., Lawrence, G.J., Finnegan, E.J., and Anderson, P.A. (1995). Contrasting complexity of two rust resistance loci in flax. *Proc. Natl. Acad. Sci. USA* **92,** 4185-4188.

Eshed, Y., Abu-Abied, M., Saranga, Y., and Zamir, D. (1992). *Lycopersicon esculentum* lines containing small overlapping introgressions from *L. pennellii*. *Theor. Appl. Gen.* **83,** 1027-1034.

Eshed, Y., and Zamir, D. (1994). A genomic library of *lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. *Euphytica* **79,** 175-179.

Eshed Y. and Zamir, D. (1995). An introgression line population of *L. pinnellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. *Genetics*, in press.

Eyal, Y., Meller, Y., Lev-Yadun, S. and Fluhr, R. (1993) A basic-type promoter directs ethylene responsiveness and abscission zone-specific expression. *The Plant J.* **4,**225-235.

Gallie, D.R. et al. (1987) *Nucl. Acid Res.* **15,** 3257-3273.

Gordon-Kahn, W.J. et al. (1990) *Plant Cell* **2,** 603-618.

Grant, M.R., Laurence, G., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.w., Dangl, J. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* **269,** 843-846.

Greve, H.D. et al. (1983) *J. Mol. Appl. Genet.* **1,** 499-511.

Hayashimoto, A. et al. (1990) *Plant Physiol.* **93**, 857-863.

Horsch, R.B. et al., (1985) *Science* **227**, 1229-1231.

Johal, G.S., and Briggs, S.P. (1992). Reductase activity encoded by the *HM1* disease resistance gene in maize. *Science* **258,** 985-987.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G. (1994). Isolation of the tomato Cf-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266,** 789-793.

Jones, J.D.G., Shlumukov, L., Carland, F., English, J., Scofield, S.R., Bishop, G., and Harrison, K. (1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. *Transgenic Research* **1,** 283-297.

K be, B., and Deisenhofer, J. (1993). Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366,** 751-756.

Kobe, B., and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *TIBS* **19**, 415-421.

Kobe, B., and Deisenhofer, J. (1995). A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**, 183-185.

Laterrot, H., and Moretti, A. (1995). Confirmation of the linkage *Tm-2*, *ah*, *Frl*. *TGC* **45**, 29.

Lawrence, G. J., Finnegan, E. J., Ayliffe, M. A., and Ellis J. G. (1995). The *L6* gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. *Plant Cell* **7**, 1195-1206.

Martin, G. B., Ganal, M. W. and Tanksley, S. (1992). Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. Mol. Gen. Genet. **233**, 25-32.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**, 1432-1436.

Mindrinos, M., Katagiri, F., Yu, G.-L., and Ausubel, F.M. (1994). The *A. thaliana* disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine rich repeats. *Cell* **78**, 1089-1099.

Ori, N., Paran, I., Aviv, D., Eshed, Y., Tanksley, S., Zamir, D., and Fluhr, R. (1994). A genomic search for the gene conferring resistance to Fusarium wilt in tomato. *Euphytica* **79**, 201-204.

Ross, D.T., Raibaud, A., Florent, I.C., Sather, S., Gross, M.K., Storm, D.R., and Eisen, H. (1991). The trypanosome VSG expression site encodes adenylate cyclase and a leucine-rich putative regulatory gene. *EMBO J.* **10**, 2047-2053.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop a common motif in ATP- and GTP-binding proteins. *TIBS* **15**, 430-434.

Scott, J.W., and Jones, J.P. (1991). Genetic control of resistance to races 1, 2 and 3 of Fusarium wilt. *TGC* **41**, 47.

Segal, G., Sarfatti, M., Schaffer, M.A., Ori, N., Zamir, D., and Fluhr, R. (1992). Correlation of genetic and physical structure in the region surrounding the I_2 *Fusarium oxysporum* resistance locus in tomato. *Mol. Gen. Genet.* **231**, 179-185.

Smiley, B.L., Stadnyk, A.W., Myler, P.J., and Stuart, K. (1990). The trypanosome leucine repeat gene in the variant surface glycoprotein expression site encodes a putative metal-binding domain and a region resembling protein-binding domains of yeast, *Drosophila* and mammalian proteins. *Mol. Cell. Biol.* **10**, 6436-6444.

Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D.G. (1995). Molecular Genetics of Plant Disease Resistance. *Science* **268**, 661-666.

Tanksley, S.D., Ganal, M.W., Prince, J.C., de Vicente, M.C., Bonierabale, M.W., Broun, P., Fulton, T.M., Giovanonni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W. and Young, N.D. (1992). High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**: 1141-1160.

Tanksley, S.D., Ganal, M.W. and Martin, G.B. (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *TIG* **11**, 63-68.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene *N*: Similarity to toll and the interleukin-1 receptor. *Cell* **78**, 1101-1115.

Young, N.D., Zamir, D., Ganal, M.W., and Tanksley, S.D. (1988). Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. *Genetics* **120**, 579-585.

CLAIMS:

1. A DNA molecule selected from the group comprising:

(i) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-1 (SEQ. ID. NO.:1);

(ii) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-2 (SEQ. ID. NO.:2);

(iii) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-3 (SEQ. ID. NO.:3);

(iv) a DNA molecule having a nucleotide sequence derived from the coding region of the clone herein designated I2C-4 (SEQ. ID. NO.:4);

(v) a DN A molecule characterized by containing a coding sequence representing at least 60% similarity with the encoded open reading frame in the DNA sequence of at least one of the DNA molecules (i) and (ii);

(vi) a DNA molecule capable of hybridization with any one of the DNA molecules (i)- (v) under moderately stringent conditions;

(vii) a DNA molecule that differs, by insertion, deletion or as a result of the degenerative nature of the genetic code, from the DNA sequences (i)-(vi); and

(viii) a fragment of any of the DNA molecules (i)- (vii).

2. A DNA molecule according to claim 1 which is expressed in tomato plants and confers resistance to *Fusarium oxysporum* f.sp *lycopersici* race 2 in said tomato plants.

3. A gene construct containing a DNA molecule according to claim 2.

4. A gene construct according to claim 3 wherein said DNA molecule includes regulatory sequences that flank the coding region.

5. A cosmid into which is subcloned a gene construct according to claim 4.

6. A gene construct according to claim 3 which further comprises DNA regulatory sequences enabling the expression of said DNA molecule in plant cells.

7. An expression vector comprising a gene construct according to claim 6 wherein said DNA regulatory sequences comprise a plant promoter, a DNA sequence that enhances translation of the mRNA transcribed from said DNA molecule and a polyadenylation/terminator sequence.

8. A tomato cell line or plant line transformed with a cosmid according to claim 5 or with an expression vector according to claim 6 or 7.

9. Tomato plants regenerated from transformed cells according to claim 8.

10. A method of selective breeding of Fusarium resistant tomatoes employing a DNA molecule according to claim 1 as a direct restriction fragment length polymorphism (RFLP) probe, which comprises:

(i) marking said DNA molecule with a suitable marker; and

(ii) reacting said probe of (i) with DNA extract of a tomato plant under hybridization conditions;

thus obtaining restriction-length polymorphism that is indicative of a resistance-type gene, which facilitates the selection of progeny that contains said resistance gene.

11. A method of screening new diseases in plants of the Solanaceae family employing a DNA molecule according to claim 1 as a direct RFLP probe which comprises:

(i) marking said DNA molecule with a suitable marker; and

(ii) reacting said probe of (i) with DNA extracts from said Solanaceae species,

thus identifying the homologous gene family in those species which can be linked to known resistance genes in those species.

12. Use of a DNA molecule according to claim 1 as a direct RFLP probe for selective breeding of Fusarium resistant tomatoes.

13. Use of a DNA molecule according to claim 1 as a direct RFLP probe for screening new diseases in plants of the Solanaceae family.

FIG. 1A

FIG. 1B

FIG. 1C

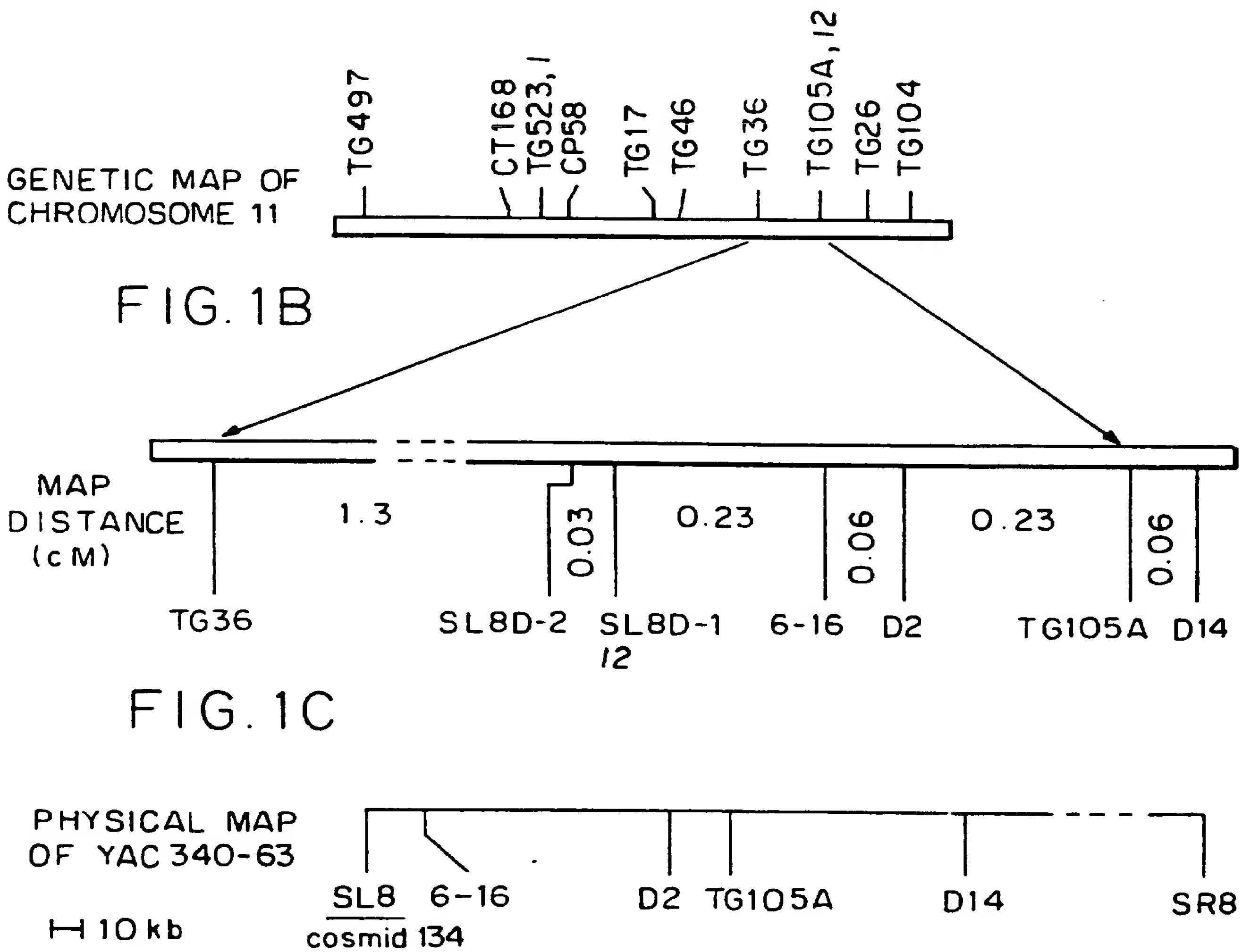


FIG. 2

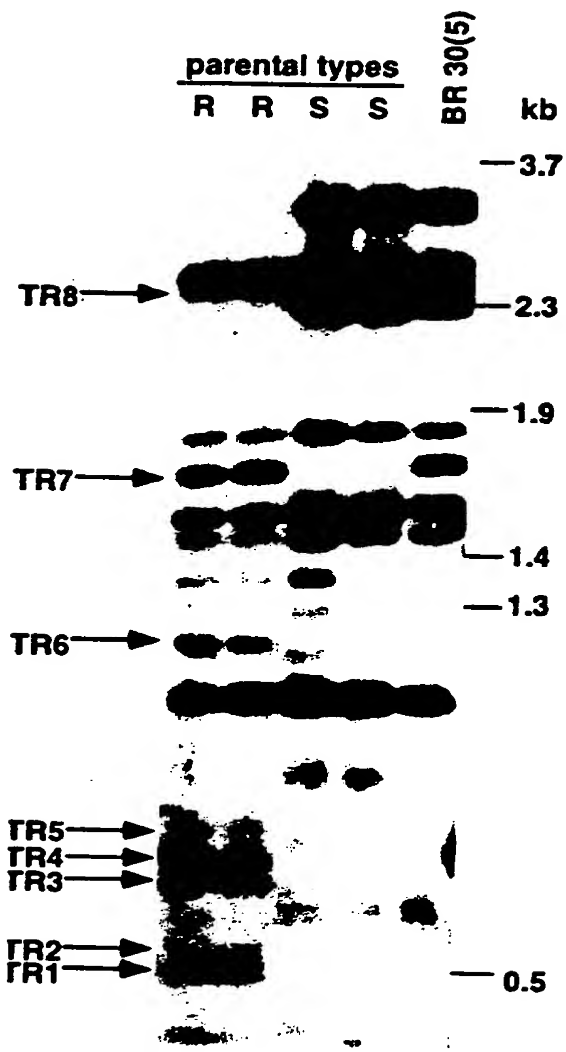

parental types
R R S S
BR 30(5)
kb
3.7
TR8
2.3
1.9
TR7
1.4
1.3
TR6
TR5
TR4
TR3
TR2
TR1
0.5

FIG. 3A

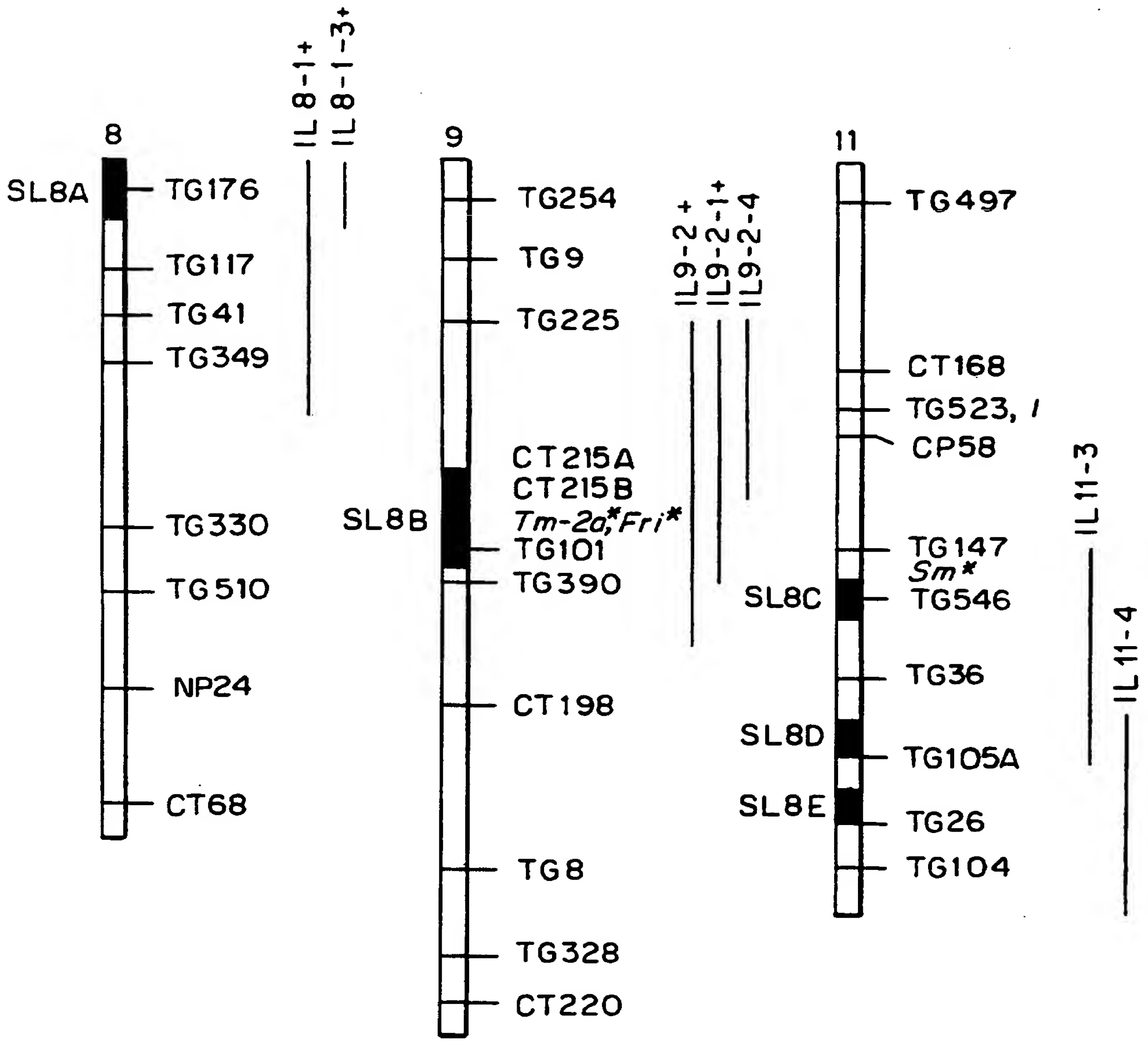

8
SL8A
TG176
TG117
TG41
TG349
TG330
TG510
NP24
CT68
IL8-1+
IL8-1-3+
9
TG254
TG9
TG225
CT215A
CT215B
SL8B
Tm-2a*Fri*
TG101
TG390
CT198
TG8
TG328
CT220
IL9-2+
IL9-2-1+
IL9-2-4
11
TG497
CT168
TG523, I
CP58
TG147
Sm*
SL8C
TG546
TG36
SL8D
TG105A
SL8E
TG26
TG104
IL11-3
IL11-4

FIG. 3B

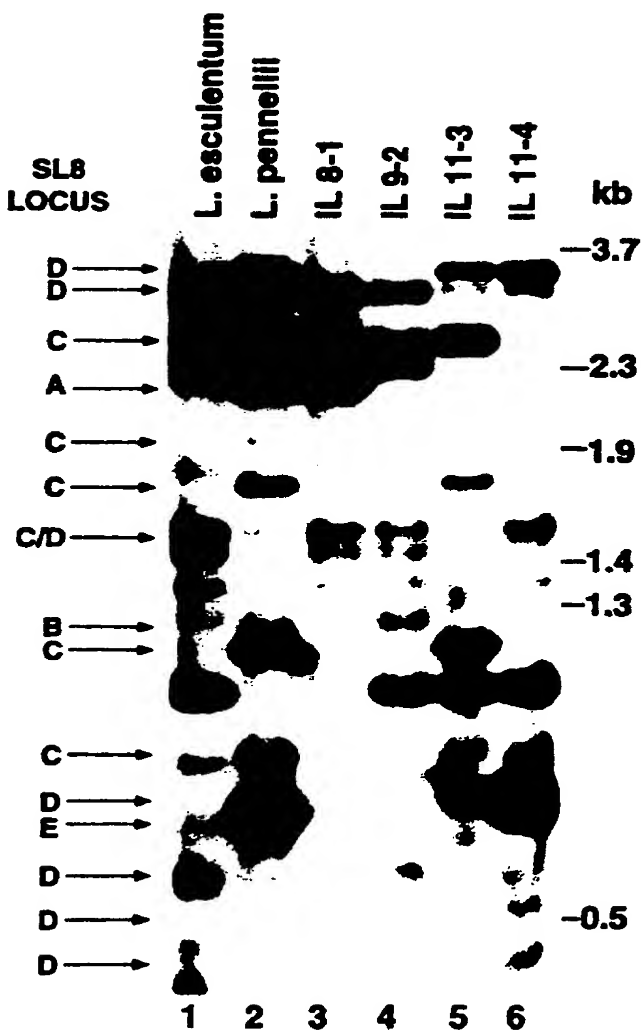

SL8 LOCUS
L. esculentum
L. pennellii
IL 8-1
IL 9-2
IL 11-3
IL 11-4
kb
D
D
C
A
C
C
C/D
B
C
C
D
E
D
D
D
3.7
2.3
1.9
1.4
1.3
0.5
1 2 3 4 5 6

FIG. 4A

```
-298 ATCTCAACTCTTCACAATTCTAAAATGTATTGATTATATGGTAGGCCCCACTACAATGAGCTGTTGAAAAATGGACTCAATTATTCTATTCAAGTCAATACTTTTTGAGA
-188 AGAGAAAAGAGGACAGTTTTGATTAGCTGAAAAATAGTTGTACTAATTTCAAGCAATCACTAATCATTCTTGTTTGCAAATTACTCTCAACATTCAACGGGTAAAGTGAA
- 78 GTACTATCTTTGTAGTTGAAGATAGAwAGAAAAATTTATCTTCTCAAAATCATTTTGTGTTCCCTTGCAGATTTGAGAAATGGAGATTGGGTTAGCAATTGGTGGTGCAT
                                                                                                  M  E  I  G  L  A  I  G  G  A  F
  32 TTCTCTCCTCAGCTTTGAATGTTCTGTTTGATAGGCTTGCTCCTAATGGTGATCTGCTCAACATGTTTCGGAAGCATACAGATGATGTTGAGCTCTTTGAGAAGCTGGGG
      L  S  S  A  L  N  V  L  F  D  R  L  A  P  N  G  D  L  L  N  M  F  R  K  H  T  D  D  V  E  L  F  E  K  L  G
 142 GACATTTTGCTTAGTCTTCAAATTGTGCTAAGTGATGCAGAGAATAAGAAAGCATCGAATCAATTTGTGAGCCAGTGGTTACATAAGCTTCAGACTGCTGTGGACGCTGC
      D  I  L  L  S  L  Q  I  V  L  S  D  A  E  N  K  K  A  S  N  Q  F  V  S  Q  W  L  H  K  L  Q  T  A  V  D  A  A
 252 TGAAAACTTGATAGAACAAGTCAATTATGAAGCTTTGAGGCTTAAAGTGGAAACAAGCAACCAGCAAGTAAGTGACCTCAACCTGTGCTTGAGTGATGATTTCTTTCTTA
      E  N  L  I  E  Q  V  N  Y  E  A  L  R  L  K  V  E  T  S  N  Q  Q  V  S  D  L  N  L  C  L  S  D  D  F  F  L  N
 362 ACATAAAGAAGAAGTTGGAAGACACTATTAAAAAACTGGAGGTGTTGGAAAAGCAAATTGGTCGCCTTGGCTTAAAGGAGCATTTTATTTCGACCAAACAAGAAACTAGA
      I  K  K  K  L  E  D  T  I  K  K  L  E  V  L  E  K  Q  I  G  R  L  G  L  K  E  H  F  I  S  T  K  Q  E  T  R
 472 ACACCTTCAACTTCTTTGGTTGATGATTCTGGTATCTTTGGAAGGAAGAATGAAATAGAGAATTTGGTTGGCCGTTTGTTGTCTATGGATACAAAGCGAAAAAATCTGGC
      T  P  S  T  S  L  V  D  D  S  G  I  F  G  R  K  N  E  I  E  N  L  V  G  R  L  L  S  M  D  T  K  R  K  N  L  A
 582 TGTAGTTCCTATTGTGGGAATGGGCGGCATGGGTAAGACAACACTTGCTAAAGCCGTTTACAATGATGAGAGAGTGCAGAAACATTTTGGTTTGACAGCTTGGTTTTGTG
      V  V  P  I  V  G  M  G  G  M  G  K  T  T  L  A  K  A  V  Y  N  D  E  R  V  Q  K  H  F  G  L  T  A  W  F  C  V
 692 TTTCTGAGGCATATGATGCTTTCAGAATACCAAAAGGTTTACTTCAAGAAATTGGATCAACTGACTTGAAGGCTGATGACAATCTTAATCAGCTACAAGTCAAATTGAAG
      S  E  A  Y  D  A  F  R  I  P  K  G  L  L  Q  E  I  G  S  T  D  L  K  A  D  D  N  L  N  Q  L  Q  V  K  L  K
 802 GCTGATGACAATCTTAATCAGCTACAAGTCAAATTGAAGGAAAAGCTGAATGGAAAAAGGTTTCTTGTTGTCCTTGATGACGTGTGGAATGATAATTATCCTGAGTGGGA
      A  D  D  N  L  N  Q  L  Q  V  K  L  K  E  K  L  N  G  K  R  F  L  V  V  L  D  D  V  W  N  D  N  Y  P  E  W  D
 912 TGACTTGAGAAATCTTTTTTTACAAGGGGATATAGGAAGTAAGATCATTGTAACGACACGTAAAGAGAGTGTTGCCTTGATGATGGATAGTGGGGCAATCTACATGGGAA
      D  L  R  N  L  F  L  Q  G  D  I  G  S  K  I  I  V  T  T  R  K  E  S  V  A  L  M  M  D  S  G  A  I  Y  M  G  I
1022 TTCTGTCTAGTGAAGACTCTTGGGCTCTATTCAAACGACATTCATTAGAGCACAAGGATCCCAAGGAACATCCAGAATTTGAAGAGGTTGGAAAACAAATTGCAGACAAG
      L  S  S  E  D  S  W  A  L  F  K  R  H  S  L  E  H  K  D  P  K  E  H  Q  E  F  E  E  V  G  K  Q  I  A  D  K
1132 TGCAAAGGGTTGCCTTTAGCTCTAAAAGCACTTGCTGGTATGTTACGCAGCAAATCAGAGGTGGATGAGTGGAGAAACATTTTACGAAGTGAAATATGGGAGCTTCCAAG
      C  K  G  L  P  L  A  L  K  A  L  A  G  M  L  R  S  K  S  E  V  D  E  W  R  N  I  L  R  S  E  I  W  E  L  P  S
1242 TTGTTCGAATGGTATATTACCAGCGCTAATGTTGAGCTACAATGATCTCCCTGCACATTTAAAGCAATGTTTGGCTTATTGTGCAATATATCCCAAAGATTATCAATTTC
      C  S  N  G  I  L  P  A  L  M  L  S  Y  N  D  L  P  A  H  L  K  Q  C  L  A  Y  C  A  I  Y  P  K  D  Y  Q  F  R
1352 GCAAAGAGCAAGTTATTCACCTGTGGATTGCTAATGGTCTTGTACATCAGTTTCATTCGGGTAACCAATACTTTATCGAGTTGAGATCAAGATCATTGTTCGAAATGGCC
      K  E  Q  V  I  H  L  W  I  A  N  G  L  V  H  Q  F  H  S  G  N  Q  Y  F  I  E  L  R  S  R  S  L  F  E  M  A
1462 TCAGAGCCTTCTGAAAGAGACGTAGAGGAATTCTTAATGCATGACCTTGTCAATGATTTGGCACAAATTGCATCTTCAAATCATTGTATAAGGTTGGAAGATAACAAAGG
      S  E  P  S  E  R  D  V  E  E  F  L  M  H  D  L  V  N  D  L  A  Q  I  A  S  S  N  H  C  I  R  L  E  D  N  K  G
1572 ATCGCATATGTTGGAACAATGTCGGCACATGTCCTATTCAATAGGACAAGATGGTGAGTTTGAGAAATTGAAATCACTCTTTAAATCAGAGCAGCTGAGGACATTACTTC
      S  H  M  L  E  Q  C  R  H  M  S  Y  S  I  G  Q  D  G  E  F  E  K  L  K  S  L  F  K  S  E  Q  L  R  T  L  L  P
```

FIG. 4B

```
1682 CAATCGATATCCAGTTCCATTACTCAAAAAAACTAAGCAAGAGGGTGTTGCATAACATACTGCCTACACTAAGATCCTTGAGGGCACTATCATTGTCTCATTACCAGATT
      I  D  I  Q  F  H  Y  S  K  K  L  S  K  R  V  L  H  N  I  L  P  T  L  R  S  L  R  A  L  S  L  S  H  Y  Q  I
1792 GAGGTGTTGCCAAATGACTTGTTTATCAAATTAAAGCTCCTCAGATTTTTGGACCTTTCTGAGACATCGATTACAAAGTTGCCGGATTCCATTTTTGTGTTGTATAACTT
     E  V  L  P  N  D  L  F  I  K  L  K  L  L  R  F  L  D  L  S  E  T  S  I  T  K  L..R..D..S..I..F..V..L..Y..N..L
1902 AGAGACACTTCTCCTGTCATCTTGTGAATATCTTGAGGAGCTACCGCTGCAGATGGAGAAGTTGATTAACTTGCGTCATCTTGACATAAGCAACACTCGGCGCTTGAAGA
     .E..T..L..L..L..S..S..C..E..Y..L  E  E  L  P  L  Q  M  E  K  L  I  N  L  R  H  L  D  I  S  N  T  R  R  L  K  I
2012 TcCCACTACATCTGAGCAGGTTGAAAAGCCTCCAAGTGTTGGTGGGAGCCAAGTTTCTTGTAGGTGGTTGGAGAATGGAATATTTGGGTGAAGCACCCAACTTATATGGA
      P  L  H  L  S  R  L  K  S  L  Q  V  L  V  G  A  K  F  L  V  G  G  W  R  M  E  Y  L  G  E  A  P  N  L  Y  G
2122 TCTCTATCAATTCTAGAGTTGGAAAATGTGGTTGATAGAAGGGAAGCTGTGAAGGCAAAGATGAGGGAGAAGAATCATGTTGAGCAATTATCATTGGAGTGGAGTGAAAG
     S  L  S  I  L  E  L  E  N  V  V  D  R  R  E  A  V  K  A  K  M  R  E  K  N  H  V  E  Q  L  S  L  E  W  S  E  S
2232 CATTAGTGCTGACAATTCACAAACAGAAAGAGACATACTTGATGAGCTACGCCCACATAAAAACATTAAAGCAGTTGAAATCACTGGATATAGAGGGACAAACTTTCCAA
      I  S  A  D  N  S  Q  T  E  R  D  I  L  D  E  L  R  P  H  K  N  I  K  A  V  E  I  T  G  Y  R  G  T  N  F  P  N
2342 ACTGGGTAGCTGATCCTTTGTTTGTTAAGCTGGTGCATTTGTATCTTAGAAACTGCAAGGACTGTTACTCCTTGCCAGCACTAGGACAACTCCCTTGTTTGGAATTCCTT
      W  V  A  D  P  L  F  V  K  L  V  H  L  Y  L  R  N  C  K  D  C  Y  S  L  P  A  L  G  Q  L  P  C  L  E  F  L
2452 TCCATTAGAGGGATGCATGGGATAAGAGTGGTGACAGAAGAGTTCTATGGCAGATTGTCCTCCAAAAAGCCTTTTAACTCTCTTGTGAAGCTTAGATTTGAAGATATGCC
     S  I  R  G  M  H  G  I  R  V  V  T  E  E  E  Y  G  R  L  S  S  K  K  P  F  N  S  L  V  K  L  R  F  E  D  M  P
2562 TGAATGGAAGCAATGGCACACACTAGGAATTGGAGAGTTCCCTACACTTGAGAAACTTTCCATTAAAAATTGCCCTGAGCTCAGTTTGGAGATACCCATCCAATTTTCAA
      E  W  K  Q  W  H  T  L  G  I  G  E  F  Q  T  L  E  K  L  S  I  K  N  C  P  E  L  S  L  E  I  P  I  Q  F  S  S
2672 GTTTAAAAAGGTTAGATATATGTGATTGTAAGTCTGTTACCTCCTTTCCTTTTAGCATACTGCCAACTACCTTGAAGAGAATAAAGATATCTGGTTGcCCAAAAATTGAAA
      L  K  R  L  D  I  C  D  C  K  S  V  T  S  F  Q  F  S  I  L  P  T  T  L  K  R  I  K  I  S  G  C  P  K  L  K
2782 TTGGAGGCGCCAGTTGGTGAGATGTTTGTGGAGTATTTGAGTGTGATTGATTGtGGTTGTGtAGATGAtATATCACCTGAGTTTCTCCCAACAGCACGTCAATTGAGTAT
     L  E  A  P  V  G  E  M  F  V  E  Y  L  S  V  I  D  C  G  C  V  D  D  I  S  P  E  F  L  P  T  A  R  Q  L  S  I
2892 TGAGAATTGCCACAACGTTACTAGGTTTTTGATTCCTACTGCCACTGAAAGTCTCCATATTCGGAATTGTGAAAAACTCTCGATGGCATGTGGAGGAGCGGCCCAGCTGA
     E  N  C  H  N  V  T  R  F  L  I  P  T  A  T  E  S  L  H  I  R  N  C  E  K  L  S  M  A  C  G  G  A  A  Q  L  T
3002 CGTCACTGAATATTTGGGGATGTAAGAAGCTCAAGTGTCTTCCAGAACTCCTTCCaTCTCTCAAGGAACTGCGACTGACTTATTGTCCAGAAATAGAAGGAGAATTGCCC
      S  L  N  I  W  G  C  K  K  L  K  C  L  P  E  L  L  P  S  L  K  E  L  R  L  T  Y  C  P  E  I  E  G  E  L  P
3112 TTCAATTTACAAATACTCGATATCAGATATTGCAAGAAACTGGTGAATGGCCGAAAGGAGTGGCATTTACAGAGACTCACAGAGTTATGGATCAAACATGATGGGAGTGA
     F  N  L  Q  I  L  D  I  R  Y  C  K  K  L  V  N  G  R  K  E  W  H  L  Q  R  L  T  E  L  W  I  K  H  D  G  S  D
3222 CGAACATATTGAACATTGGGAGTTGCCTTCCTCTATTCAGAGACTATTCATATTCAATCTGAAAACATTAAGCAGCCAACATCTCAAAAGCCTCACCTCTCTTCAATTTC
      E  H  I  E  H  W  E  L  P  S  S  I  Q  R  L  F  I  F  N  L  K  T  L  S  S  Q  H  L  K  S  L  T  S  L  Q  F  L
3332 TACGTATTGTTGGTAATTTATCTCAGTTTCAGTCACAAGGCCAACTTTCCTCCTTTTCTCACCTCACTTCGCTTCAAACTCTACAAATCTGGAATTTTCTTAATCTTCAA
      R  I  V  G  N  L  S  Q  F  Q  S  Q  G  Q  L  S  S  F  S  H  L  T  S  L  Q  T  L  Q  I  W  N  E  L  N  L  Q
```

7 / 18

3442 TCACTACCTGAATCAGCACTGCCCTCCTCCCTCTCTCACCTGATCATCTCCAATTGCCCTAATCTCCAATCCCTTCCATTAAAAGGGATGCCCTCTTCCCTCTCTACACT
S L P E S A L P S S L S H L I I S N C P N L Q S L P L K G M P S S L S T L
3552 ATCAATTTCCAAATGTCCATTGCTCACACCACTACTAGAATTTGACAAGGGGGAATACTGGACAGAAATTGCTCATATCCCCACCATACAGATCGATGAGGAATGCATGT
S I S K C P L L T P L L E F D K G E Y W T E I A H I P T I Q I D E E C M
3662 AATGATTAAAACAAATGGCTCCCCAACTGATGTAAGCTATTCTTTTCCCTCATAAGCTTTTTATTTCACTTTGCTTTTGGGGTTACTTCTTTTCATTTTTAATCATGCCG
3772 GGCTAGCTCATCATCAAACACATAGCATTATATTTAACCTCCATAGAGAATCTAAATTTTTTAAAGGATAATTCAATCACAAGTTTTAGGAAATAACTGCAACTTCCATT
3882 GTCACATGTTATATGATTCTATGTTTCTCATGGCTTATTGGTTTATGCTCTTACCGTGTTTGAATTCACGTCTCAATTGCCACCATGTTTAATCAAAAGTTTTTAGTTCT
3992 TGTAATCATCAACCATCCTATGTCACTAGCAATTTGGATAGATAAAAGAGGTAGACAAAAAAAGCGAAACATCTTTTTTCTTTCGTATAGCGACCAGACGACTACATTTT
4102 GATAGGTAAGGGCTATAGATATACATTTGCAGGGTGTTAATCCAACGAGTTAGAAAATCCCTGTCTTTAGATATCTTCTCTTGCATATACTTTGGCAATTTTAAGCTACA
4212 GTTTGAACTCATGTGTTGTTGCTAACTTAAACATGTTTTGTGCTTAATCAGATGTGGATTTTGAAGAGCGAGTACGACAAGTCTGGTACATTAATTGTCCGTAGGAAGTG
4322 TTTCTAAGGTGCTGCTGCTATTTTTACATCTGTTCCCGAGTTTGGTTTTTTTTTAAATCTTTCCACTAAAGCTATTATGTCGTCCACAGTGAATTTTCAGGTCTGTTGT
4432 TATAGGCAAGTCTTTGAGATGCGACTATCAAAGAAGGGCGATTACAATCAGTGTACCGCTGAAACTATTTCATGTTTCCAGTGCAAGCCTCTTTTGTAAGTTGACAAACT
4542 CGATTAGTTAATATGTTTGGGACTCAACTAGTGGTTAGAGTACTCATTTTGTAAGACTTGgGTACAGAAAATCAAATTAGAATTATCACTCGCGATGGTTGGAATAAC

FIG. 4C

FIG.5A

```
-202 AAGAGGGAAAGTTTTGCTTCTTGTAGATTAGCTAGAACATAGTTGTACTAAACAAAGAGAGAGCAATCGCTGATCATTCTTGTTTGCAAATTACTCTTAGCATTCGACAG
 -92 GTAAAATAAAGCTCATAACAGACACTATTTGAATATATATTTTGTTAAAGAATCATATTGTGTTTCCTTGTTTTGCTTTTGCAGATTTGAGAAATGGAGATTGGCTTAGC
 -30                                                                                                 M E I G L A
  18 ATTTGGTGGTGCATTTCTCTCCTCAGCTTTGAATGTTCTGTTTGATAGGCTTGCTCCTAACGGTGATCTGCTCAACATGTTTCGGAAGCATAAGGATCATGTTAAGCTCT
   7 V G G A F L S S A L N V L F D R L A P N G D L L N M F R K H K D H V K L L
 128 TAAAGAAGCTGAAAATGACTTTGCGTGGTATTCAGATTGTGCTAAGTGATGCAGAGAATAAGCAAGCATCAAATCCATCTGTGAGAGACTGGCTTAATGAGCTTCGAGAT
  44 K K L K M T L R G I Q I V L S D A E N K Q A S N P S V R D W L N E L R D
 238 GCTGTCGACTCTGCTGAAAATTTAATAGAAGAAGTCAATTATGAAGCTTTGAGGCTTAAGGTGGAAGGTCAGCATCAGAACTTTTCAGAAACAAGCAACCAGCAAGTAAG
  80 A V D S A E N L I E E V N Y E A L R L K V E G Q H Q N F S E T S N Q Q V S
 348 TGATGAATTTTTCCTTAACATAAAGGACAAGTTGGAAGACACTATTGAAACATTAAAGGATTTGCAAGAGCAAATTGGTCTCCTTGGCTTAAAGGAGTATTTTGATTCCA
 117 D E F F L N I K D K L E D T I E T L K D L Q E Q I G L L G L K E Y F D S T
 458 CGAAACTAGAAACTAGAACACCTTCAACTTCTTTGATTGATGAACCAGATATCTTTGGTAGGCAGAGCGAAATAGAGGATTTGATTGACCGTCTATTGTCTGAAGGTGCA
 154 K L E T R T P S T S L I D E P D I F G R Q S E I E D L I D R L L S E G A
 568 AGTGGGAAAAATCTGACAGTGGTTCCTATTGTTGGAATGGGTGGCTTGGGCAAGACAACACTTGCTAAAGCCGTATACAATGATGAGAGTGTGAAGAACCATTTTGATTT
 190 S G K N L T V V P I V G M G G L G K T T L A K A V Y N D E S V K N H F D L
 678 GAAAGCTTGGTTTTGTGTTTCCGAGGCGTATAATGCTTTCAGAATAACAAAAGGGTTACTTCAAGAAATTGGCTCAATTGACTTAGTTGATGACAATCTTAATCAGCTAC
 227 K A W F C V S E A Y N A F R I T K G L L Q E I G S I D L V D D N L N Q L Q
 788 AAGTCAAATTGAAGGAAAGATTAAAGGAAAAGAAGTTTCTTATCGTTCTGGATGATGTGTGGAATGACAACTACAACGAGTGGGATGAATTGAGAAATGTTTTTGTACAA
 264 V K L K E R L K E K K F L I V L D D V W N D N Y N E W D E L R N V F V Q
 898 GGAGATATAGGAAGTAAGATCATTGTGACGACACGCAAAGACAGTGTTGCCTTGATGATGGGAAATGAGCAAATTAGCATGGGCAATTTGTCTACCGAAGCCTCTTGGTC
 300 G D I G S K I I V T T R K D S V A L M M G N E Q I S M G N L S T E A S W S
1008 TTTATTTCAAAGACATGCATTTGAAAACATGGATCCTATGGGACATTCGGAACTTGAAGAGGTTGGAAGACAAATTGCAGCTAAGTGCAAAGGACTGCCCTTAGCTCTGA
 337 L F Q R H A F E N M D P M G H S E L E E V G R Q I A A K C K G L P L A L K
1118 AGACGCTTGCTGGCATGTTACGCTCCAAATCAGAGGTTGAAGAGTGGAAATGTATTCTGAGAAGTGAAATATGGGAGCTGCGAGACAATGACATATTACCAGCGTTAATG
 374 T L A G M L R S K S E V E E W K C I L R S E I W E L R D N D I L P A L M
1228 TTGAGCTACAATGATCTTCCTGCACATTTAAAGCGATGCTTTTCTTTTTGTGCAATATTTCCTAAAGATTATCCATTTAGGAAAGAACAAGTTATTCATCTATGGATTGC
 410 L S Y N D L P A H L K R C F S F C A I F P K D Y P F R K E Q V I H L W I A
1338 CAATGGTCTTGTACCTGTGGAAGATGAAATAATTCAAGATTTAGGCAACCAATTCTTTCTCGAGTTGAGTTCAAGATCATTATTTGAAAGGGTCCCAAATCCTTCTGAAG
 447 N G L V P V E D E I I Q D L G N Q F F L E L S S R S L F E R V P N P S E G
1448 GAAACATAAAGGAATTATTCCTAATGCATGACCTTGTCAATGATTTAGCCCAACTTGCATCTTCAAAACTTTGTATCAGGTTGGAAGAGAGCCAAGGATCTCATATGTTG
 484 N I K E L F L M H D L V N D L A Q L A S S K L C I R L E E S Q G S H M L
1558 GAACAATGTCGGCACTTATCATATTCTATGGGATATGACGGTGGGTTTGAGAAATTGACACCCCTCTACAAATTGGAGCAGCTGAGGACATTGCTTCCGACATGTAGTAG
 520 E Q C R H L S Y S M G Y D G G F E K L T P L Y K L E Q L R T L L P T C S S
1668 TGTCAATTATTTCTATAACCCTCTAACCAAGAGGGTGTTGCATAACATACTGCCTACACTAAGATCCTTAAGGGCATTATCATTGTCTCATTACAAGATGGAGGAGTTGC
 557 V N Y F Y N P L T K R V L H N I L P T L R S L R A L S L S H Y K M E E L P
1778 CAAATGACTTGTTTATCAAATTAAAGCTCCTCAGATTTTTGGATATTTCTCGGACAAATATTAAAAGGTTGCCAGATTCCATTTGTGTGTTGTATAACTTGGAGACACTT
 594 N D L F I K L K L L R F L D I S R T N I K R L P D S I C V L Y N L E T L
1888 CTCCTTTCATCTTGTAAACTTGAGGAGCTACCGCTGCAGATGGAGAAGTTGATTAACTTGCGTCATCTTGACATAAGCAACACTTGGCACTTGAAGATGCCACTACATCT
 630 L L S S C K L E E L P L Q M E K L I N L R H L D I S N T W H L K M P L H L
```

FIG. 5B

1998 GAGCAGGTTGAAAAGCCTCCAAGTGTTGGTGGGAGCCAAGTTTCTTGTAGGTGTTTGGAGAATGGAAGATTTGGGTGAAGCACAAAACTTATATGGATCTCTATCAGTTG
667 S R L K S L Q V L V G A K F L V G V W R M E D L G E A Q N L Y G S L S V V
2108 TAAAGTTGGAAAATGTGGTTGATAGAAGGGAAGCTGTGAAGCCAAAGATGAGGGAGAAGAATCATGTTGAGCAATTATCATTGGAGTGGAGTGAAAGCATTAGTGCTGAC
704 K L E N V V D R R E A V K P K M R E K N H V E Q L S L E W S E S I S A D
2218 AATTCACAAACAGAAAGAGACATACTTGATGAGCTACGCCCmCATAAAAATATTCAAGAAGTCAAAATCATTGGATATAGAGGGACAAACTTTCCCAATTGGGTAGCTGA
740 N S Q T E R D I L D E L R P H K N I Q E V K I I G Y R G T N F P N W V A D
2328 TCCTTTGTTTCTTAAGCTGGTGAAATTGTCTCTTAGAAACTGCAAGGACTGTTACTCCTTGCCAGCACTAGGACAACTCCCTTGTTTGAAATTCCTTTCCGTTAAAGGGA
777 P L F L K L V K L S L R N C K D C Y S L P A L G Q L P C L K F L S V K G M
2438 TGCATGGAATAAGAGTGGTGACGGAAGAATTCTACGGCAGATTGTCCTCCAAAAAGCCTTTTAACTGTCTTGAGAAGCTTGAATtTGAAGATATGACAGAGTGGAAGCAA
814 H G I R V V T E E F Y G R L S S K K P F N C L E K L E F E D M T E W K Q
2548 TGGCACGCACTAGGAATTGGAGAGTTCCCTACACTTGAGAAGCTTTCAATTATAAATTGCCCTGAGCTCAGTTTGGAGATACCGATCCAATTTTCAAGTTTAAAAAGGTT
850 W H A L G I G E F P T L E K L S I I N C P E L S L E I P I Q F S S L K R F
2658 TAGGGTTTTTGGTTGTCCAGTTGTTTTTATGATGCTCAAGTGTTAAGATCCCAACTTGAGGGAATGAAGCAGATTGAGGAAATATATATCCGTGATTGTAACTCTGTTA
887 R V F G C P V V F Y D A Q V L R S Q L E G M K Q I E E I Y I R D C N S V T
2768 CCTCCTTTCCTTTTAGCATACTGCCAACTACCTTGAAGACAATAGACATATCTGGTTGCCCAAAATTGAAATTGGAGGCGCCAGTTTGTGAGATGAGTATGTTTCTGGAG
924 S F P F S I L P T T L K T I D I S G C P K L K L E A P V C E M S M F L E
2878 GAATTTAGTGTGGAAGAATGTGGTTGTGTATCACCTGAGTTTCTCCCAACAGCACGTGAATTGAGAATTGGGAATTGCCACAACGTAAGGTTTTTGATTCCTACTGCCAC
960 E F S V E E C G C V S P E F L P T A R E L R I G N C H N V R F L I P T A T
2988 TGAAACTCTCCATATTCGGAATTGTGAGAATGTTGAAAAACTATCGATGGCATGTGGAGGAGCTgCCCAGCTGACATCACTGGATATTTCGGGATGTaAGAAGCTCAAGT
997 E T L H I R N C E N V E K L S M A C G G A A Q L T S L D I S G C K K L K C
3098 GTCTTCCAGAACTCCTTCCATCTCTCAAGGAACTGCAACTGACTAATTGTCCAGAAATAGAAGGAGAATTACCCTTCAATTTACAAAAACTCTATATCAGAGATTGCAAA
1034 L P E L L P S L K E L Q L T N C P E I E G E L P F N L Q K L Y I R D C K
3208 AAACTGGTGAATGGCCGAAAGGAGTGGCATTTACAGAGACTCACAAAGTTAGTGATCTATCATGATGGGAGTGACGAAGATATTGAACATTGGGAGTTGCCTTGTTCTAT
1070 K L V N G R K E W H L Q R L T K L V I Y H D G S D E D I E H W E L P C S I
3318 TACAAGACTTGAGGTATTTAATCTGATAACATTAAGCAGCCAACATCTCAAAAGCCTCACCTCTCTTCAATATCTATGTATTGATGGTAATTTATCTCCGATTCAGTCAC
1107 T R L E V F N L I T L S S Q H L K S L T S L Q Y L C I D G N L S P I Q S Q
3428 AAGGCCAGATTTCCTCCTTTTCTCACCTCACTTCGCTTCaAACTCTACAAATCTGGAATTTCCATAATCTCCAATCACTTTCAGAATCAGCACTGCCCTCCTCCCTCTCT
1144 G Q I S S F S H L T S L Q T L Q I W N F H N L Q S L S E S A L P S S L S
3538 CAGCTGGAGATCTTCCATTGCCCTAATCTCCAATCCCTTCCATTAAACGGGATGCCCTCTTCCCTCTCTAAACTATTGATTTCAGGATGTCCATTGCTCACACCACTACT
1180 Q L E I F H C P N L Q S L P L N G M P S S L S K L L I S G C P L L T P L L
3648 AGAATTTGACAAGGGGGAATACTGGCCACAAATTGCTCATATCCCCACCATACTGATCGATTGGGAATATATTTAACAATTAAAACAAATGGCTCTCCAACTGATGTAAG
1217 E F D K G E Y W P Q I A H I P T I L I D W E Y I
3758 CTATTTGTTACCCTCAGAAGCTTTTTATTTCACTTTGCTTTTTTCTTAATTCTTTTCATTTTAAATCATGTCGTGCTCATCATCAAACACATAGCTTTATAGGTAAGCTC
3868 CATACAGAATCTAACTTTTTGAAAGGATAATTCGATCACAAGTTTTAGGAAATATCTGCAACTTCCATTGTCAGATGTTATATGATTCTATGTTTCTCATGGCTTATTGG
3978 TTTATGCTCTTACCGCGTTTTCATTCACGTCTCAATTGCCAAAGAAATCATATGTAGCTTCTGCCACCATGTCCAATCGAAAGTTTTTAGTTCTTGTAATCATCAACCAT
4088 CCTATGTCACTAGAAATTTGGATAGGTAAAAGAGGTAGACAGAAAAGCTAAACATCTTTTTTCTTTCGTACAGCGACCAGACAACTACATTTGGATAGGTAAGGGCTATA
4198 GATATACATTTGCAGGGTGTTAATCTAACGAGTAAGGAAAATCACTGCCTCCAGATATCTCCTCT

10 / 18

FIG.6

1	AATTCGGCAC	GAGAATTGAA	ATTGGAGGCT	CCAGTTGGTG	AGATGTTTGT
51	GGAGTATTTG	AGTGTGAATG	ATTGTGGTTG	TGTAGAAGAT	ATATCACCTG
101	AGTTTCTCCC	AACAGCACGT	AAATTGATTA	TTACGGATTG	CCAGAACGTT
151	ACTAGGATTT	TGATTCCTAC	TGCCACTGAA	ACTCTCACTA	TTGAGAATTG
201	TGAGAATGTT	GAAAAACTAT	CGGTGGCATG	TGGAGGAGCG	GCCCAGATGA
251	CGTCTCTGAT	TATTTCGGAG	TGTAAGAAGC	TCAAGTGTCT	TCCAGAACGT
301	ATGCAGGAAC	TCCTTCCATC	TCTCAAGGAA	CTGCGTCTGT	CTGATTGTCC
351	AGAAATAGAA	GGAGAATTGC	CCTTCAATTT	ACAAAAACTC	TATATCAGTT
401	ATTGCAAGAA	ATTGGTGAAT	GGCCGAAAGG	AGTGGCATTT	ACAGAGACTC
451	ACAGAGTTAT	GGATCCATCA	TGATGGGAGT	GACGAAGATA	TTGAACATTG
501	GGAGTTGCCT	TCCTCTATTC	AGAGTCTTAC	CATATGCAAT	CTGATAACAT
551	TAAGCAGCCA	ACATCTCAAA	AGCCTCACCT	CTCTTCAATA	TCTATGTTTT
601	GATGGTAATT	TATCTCAGAT	TCAGTCACAA	GGCCAGCTTT	CCTCCTTTTC
651	TCACCTCACT	TCGCTTCAAA	CTCTACAAAT	CCGTAATCTC	CAATCACTTG
701	CTGCATTAGC	ACTGCCCTCC	TCCCTCTCTC	ACCTGACCAT	CCTCAATTTC
751	CCTAATCTCC	AATCACTTTC	AGAATCAGCA	CTGCCCTCCT	CCCTCTCTCA
801	CCTGATCATA	GATGATTGCC	CTAATCTCCA	ATCACTTTCA	GAATCAGCAC
851	TGCCCTCCTC	CCTCTCTCAC	CTGGACATCT	CCAATTGcCC	TAATCTCCAA
901	TCACTTTCAG	AATCAGCACT	GCCCTCCTCC	CTCTCTAGCC	TGACCATCTA
951	TGATTGCCCT	AATCTCCAAT	CACTtCCAGT	AAAAGGGATG	CCGTCTTCCC
1001	TCTCTGAACT	AGCAATTTCC	AAATGTCCAT	TGCTCAAACC	ACTACTAGAA
1051	TTTGGAAAGG	GGGAATACTG	GCCAAATATT	GCTCATATCC	CCTCCATATA
1101	CATCGATTGG	GAACGCATGT	AATGATTAAA	ACGAATGGCT	CCCCAACTGA
1151	TATGTGGATT	TTGAAGAGCG	AGTACGACAA	GTCTGGTACA	TCAATTGTCC
1201	GTAGGAAGTG	TTTCTAAGTG	AATTTTCAGG	TTTGTTGTTA	TAGGCAAGTC
1251	TTTGAGATGC	GACTATCAAA	GAAGGGCGAT	TACGATCAGT	GTACCGCTGA
1301	TATTATTTCA	TGTTTCCAGT	GCAAGCTTCT	TTTGTAAGTT	GACAAACTTG
1351	ATTAGTTCTC	GTGCCGAATT	C		

FIG. 7

1	CCAGGGTTTT	CCCAGTCACG	ACGTTGTAAA	ACGACGGCCA	GTGAATTGTA
51	ATACGACTCA	CTATAGGGCG	AATTGGAGCT	CCACCGCGGT	GGCGGcCGCT
101	CTAGAACTAG	TGGATcCCCC	CGGGCTGCAG	GAATTCTATG	GCAGATTGTC
151	CTCCAAAAAG	CCTTTTAACT	GTCTTGAGAA	GCTTGAATTT	GAAGATATGA
201	CGGGGTGGAA	GCAATGGCAC	GCACTAGGAA	TTGGAGAGTT	CCCTACACTT
251	GAGAACCTTT	CCATTAAAAA	TTGCCCTGAG	CTCAGTTTGA	AGATACCCAT
301	CCAATTTTCA	AGTTTAAAAA	GGTTACAAGT	TAGAGGTTGT	CCAGTTGTTT
351	TtGATGATGC	TCAACTGTTT	AGATCCCAAC	TTGAAGCAAT	GAAGCAGATT
401	GAAGCATTAt	AtATACGTGA	TTGTAACTCT	ATTACCTCCT	TTCCTTTTAG
451	CATACTGCCA	ACTACCTTGA	AGACAATAGA	GATATCTGGT	TGCCCAAAAT
501	TGAAATTCGA	GGCGCCAGTT	GGTGAGATGT	TTGTGGAGTa	TTTGAGTGtG
551	ATTGATTGTG	GTTGTGTAGA	TGATAATATC	ATTAGAGTTT	CTCCCAGCAG
601	CGTGTAAATT	GAGTATTATG	AGTTGCCACA	ACTTTACTAG	GTTTtGATT
651	CCTACTGCAA	CTGAAACTCT	CACTATTTCG	AATTGTGAGA	ATGTTGAAAA
701	ACTATCGGTG	GCATGTGGAG	GAGCGGCCCA	GATGACGTTA	CTGCATaTTT
751	tGAAGTGTAA	GAAGCTCAAG	TGTCTGCCAG	AACGTATGCA	GGAACTCCTT
801	CCATCTCTCA	AGGATTTGTA	TCTTTCCAAT	TGTCCAGAAA	TAGAAGGAGA
851	ATTGCCCTTC	AATTTACATA	AACTCCGTAT	CAGTGATTGC	AAGAAACTGG
901	TGAATGGCCG	AAAGGAGTGG	CATTTACAGA	GACTCACAGA	GTTAGTGATC
951	CATCATGATG	GGAGTGACGA	AGATATTGAA	CATTGGGAGT	TGCCTTGTTC
1001	TATTACAGAA	CTTGAGGgTA	TACAATATGA	TAACATTAAG	CAGCCAACAT
1051	CTCAAAAGCC	TCACCTCTCT	TCAATGTCTA	AGTATTGGTG	GTAATTTATC
1101	TCAGATTGGC	CGTCTTTCCT	CCTTTTCTCA	CCTCACTTCG	CTTCAAACTC
1151	TACAAATCAG	GAATTTCGGT	AATCTCCAAT	CACTTGCTGA	ATCAGCACTG
1201	CCATCCTCCC	TCTCTCACCT	GACCATCTCC	CGTTGCCCGA	ATCTCCAATC
1251	ACTTGCTGAA	TCAGCACTGC	CCTCCTCCCT	CTCTCACCTG	AACATCTATG
1301	ATTGCCCGAA	TCTCCAATTA	CTACCTGAAT	CAGCACTGCC	CTCCTCCCTC
1351	TCTCACCTGG	ACATCTCCCA	TTGTCCTAAT	CTCCAATCAC	TACCTGAATC
1401	AGCACTGCTC	TCCTCCCTCT	CTCACCTGGA	CATCTCCCAC	TGTCCTAATC
1451	TCCAATCACT	TGCTGAATCA	GCACTGCCCT	CCTCCCTCTC	TCACCTGACC
1501	ATCTCCCATT	GCCCTAATCT	CCATTCACTT	TCAGAAAAAG	GGATGCCCTC
1551	TTCCCTCTCT	AAACTATCTA	TTTCCAAATG	TTCATTGCTC	ACACCACTAC
1601	TAGAATTTAA	CAAGGGGGAA	TACTGGACAA	ATATTGCTCA	TATCTCCACC
1651	ATACAGATCG	ATTGGAAATG	CATGTAATGA	TTAAAACGAA	TGACTCCCCA
1701	ACTGATATGT	GGATTTAGAA	GAGCGAGTaC	GACAAGTCTG	GTACATCAAT
1751	TGTCCGTAGG	AAGTGTTTCT	AAGTGAATTT	TCAGGTCTGT	TGTTATAGGC
1801	AAGTCTTTGA	GATGTGACTA	TCAAAGAAGG	GCGATTACAA	TCA TGTACC
1851	GCTGATACTA	TTTCATGTTT	CCAGTGCTAC	AGTGCAAGCC	TCTTTTGTAA
1901	GTTGnCAAAC	TCGATTAGTT	AATATGTTTG	GGACTCAACT	ACTACTCATT
1951	TTGTAAGACT	TAAGTACAGA	AAATCAAATT	AGAATTATAA	CTCGCGATGG
2001	TTGAGTAAAC	TCCAAGAAGC	TCGTGCC		

12 / 18

```
I2C-1   .................................MEIGLAIGGAFLSSALNVLFDRLAPNGDLLNMFRKH 36
I2C-2   ......................................V                              36
RPM1    ......................................MAS TVDFGIGRILSV ENETL  SGVHGE 30
RPS2    ................................................................... 0
N-GENE  ..............................................MASSSSSSRWSY VFLS  GE 21
L6      MSYLREVATAVALLLPFILLNKFWRPNSKDSIVNDDDDSTSEVDAI DST PSGSFPSVEYEVFLS  GP 70
CON.

I2C-1   TDDVELFEKLGDILLSLQI.VLSDAENKKASNQFVSQWLHKLQTAVDAAENLIEQVN..YEALRLKVETS 103
I2C-2   K H K LK  KMT RGI   .         Q   PS RD  NE RD   S      E  ..       GQ 103
RPM1    I ...... MKKE  IMKS.F E T.H HGG GSTTTTTQLF  F ANTRD AY IEDILDEFGYHIHGY 92
RPS2    ......MDFISSLIVGCAQ.  CESM MAEKRGHKTD.... RQ ITDL TA GDLKAIRDD T RIQ.. 57
N-GENE  DTRKTFTSH YEV NDKG KTFQ DXRLEYGATIPGELCKAIEESQF IVVFS NYATSRWC NEL KIM 91
L6      DTREQFTDF YQS RRYK HTFR DDELLKGKEIGPNL RAIDQSKIYVPIISSGYADSKWC MELA IV 140
CON.

I2C-1   NQQVSDL.NLCLSDDFFLNIKKKLEDTIKKLEVLEKQIGRLGLXEHFIS.........TK.......... 153
I2C-2   H NF ETS QQV  E       D       ET KD QE    L      Y D .........  .......... 154
RPM1    RSCAKIWRAFHFPR..YMWARHSLAQKLGMVN MIQS SD.SM RYYH .........ENYQAALLPPID 150
RPS2    QDGLEGRSCSNRAREWLSAVQVTETK ALL VRFRRREQ TRMRRRYL CFGCADYKLC KVSAILKSIG 127
N-GENE  ECK.TRFKQTVIPIFYDVDPSHVRNQKESFAKAF EHETKYKDDVEG QRWRIALNEAANLKGSCDNRDK 160
L6      RR EE PRRII PIFYMVDPSDVRHQ GCYKKAFR HANKF..DGQT QNWKDALKKVGDLKGWHIGKND 208
CON.                                                                 I
                                                                  P-LOOP
I2C-1   ......QETRTP.STSLVDDSGIFGRKNEIEN..LVGRLLS.MDTKRKNLAVVPIVGMGGMGKTTLAKAV 213
I2C-2   ......L       .      I EPD      QS     D.. ID     .EGASG     T          L      214
RPM1    DGDAKWVNNISE. SLFFSENSLV IDAPKGK.. I      .PEPQ ...I  AV       S      SANI 213
RPS2    ELRERSEAIK D.GG IQVTCREIPI SVVG TIMMEQV E.FLSEEEERGIIGVY P  V     MQSI 195
N-GENE  TDADCIRQIVDQI SK CXI LSY.LQ IVGIDTHLEKIE LLEIGINGVRIMG W   V     I R I 229
L6      KQGAIADKVSADIWSHISKENL LETDELVG DDHITAV EKLSLDSE VTM GLY    I    T      278
CON.                                                   αα GMGGαGKTT

I2C-1   YNDERVQKH.FGLTAWFCVSEAYDAFRITKGLLQEI.GSTDLKADDNLNQLQVKLKADDNLNQLQVKLKE 281
I2C-2       S KN . D K          N              . I       ..............V          278
RPM1    FKSQS RR . ESY  VTI KS VIEDVFRTMIK FYKEA TQIPAE YS ........GYRE VE  V 274
RPS2    N ELITKG QYDVLI VQM REFGECT QQAVGARGLG W E ETGENRA ................IYR 250
N-GENE  FDTLLGR.....MDSSYQFDG CFLXD KE..NKRGNH LQNALLSE LR...EKANYN EEDGKHQMAS 289
L6        KI..........SSC.FDCCCFIDN RETQEKDGVVVLQK LVSEILRIDSGSGVGFN DSGGRKTI   337
CON.

                 II                                   III
I2C-1   KLNGKRFLVVLDDVWNDNYPEWDDLRNLFLQGDIGSKIIVTTRKESVALM...MDSGAIYMGILSSEDSW 348
I2C-2   R KE K  I             N     E    V V             D      ... GNEQ S  N  T A     335
RPM1    Y QS   YI          TTGL.. REISIALPD IY  RVMM   DMN  SFPYGIG TKHEIEL KEDEA   342
RPS2    A RQ      LL        EEIDL KTGVPR..PDRENKC VMF    .. I  CNNMGAEYKLRVEF EKKHA   315
N-GENE  R RS KV I      I.DNKDHYLEY AGGLEWFGN   R  I    DKHLI....EKNDIIYEVTA PDHE I 354
L6      RVSRFKI            .DEKFKFE M GSPKDFISQ. RF I S SMF LGTLNENQCKLYEV SM KPR L 405
CON.         K+αLαVLDDV                        S+ααTTR                          L
                                                 IV
I2C-1   ALFKRHSL.EHKDPKEHPEFEEVGKQLADKCKGLPLALKALAGMLRSKSEVDEWRNILRSEIWELPSCSN 417
I2C-2   S   Q  AF. NM  MG S L     R     A              T               E   KC             R..D 402
RPM1    V   SNKAFFPASLEQCRTQNL PIARKLVER O        IAS GS MST KFES    KKVYSTLN     NNNHE 412
RPS2    E   CSKVW..R    LL SSSIRRLAEI VS   G        IT G AMAHRETEE    IHASEVLTRFPAEMKG 384
N-GENE  Q     Q AFGKE...VPNEN    KLSLEVVNYA            VWGSL .HNLRLT   KSAIE....HMKNN.S 415
L6      E   SK AFKKN...TPPSYY TLANDVV TTA       T  VIGSL .F Q IAV EDT E....Q RRTL 467
CON.     LF                                       GLPLAL                    EW
```

FIG. 8A

```
                        V
I2C-1    ..GILPALMLSYNDLPAH.LKQCFAYCAIYPKDYQFPKEQVIHLWIANGLVHQFHS......GNQYFIEL 478
I2C-2      D              .  R  SF   F    P                  PVEDEI IQDL   F L  468
RPM1     LKIVRSIMF  F    YP.  R  L  SLF VN RMKRKRL RM M QRF EPIRGVKAEEVADS LN   481
RPS2     MNYVFAL KF  DN ESDL RS  L   LF EEHSIEI  LVEY VGE FLTSS GVNTIYK YFLIGD  454
N-GENE   YS  IDK KI  DG EPKQQEMFLDIACFLRGE...E DYILQILESCHIGAEYG...........LRI   471
L6       LDEVYDR KI  DA NPEAKEIFLDIACFFIGQ...N  EPYYM TDCNFYPASN...........I F  523
CON.        α    L αSY  L
                                      VI
I2C-1    RSRSLFEMASEPSERDVEE.FLMHDLVNDLAQ................IASSNHCIRLEDNKGSHMLEQC 531
I2C-2    S      RVPN   GNIK L               ................L    KL     ESQ        522
RPM1     VY NMLQVILWNPFGRPKA. K    VIWEI ..............................SVS 515
RPS2     KAAC L ......TG EKTQVK   NV RSF .......................LWMASEQ TYKELIL 494
N-GENE   IDK  VFI.......SEYNQVQ     IQ MGKYIVNFQKD PGERSRLWLAKEVEEVMST T TMAM AI 533
L6       IQ CMIQV.......GDDDE K    QLR MGREIVRREMVLPWKRSRIWSAEEGIDL LNK    SKVKAI 584
CON.                          α MHDαα -

I2C-1    RHMSYSIGQDGEFEKLKSLFKSEQLRTLLPI..DIQFHYSKKLSKRVLHNILPTLRSLRALSLSHYQIEV 599
I2C-2      L   M Y   G      TP Y L        TCSSVNYF .NP T                    KM E 541
RPM1     KLERFCDVYNDDSD.......GDDAAETMENYGSRHLCIQ EMTPOSIRA..TN H   LVC SAKHKM L 576
RPS2     VEP MGHTEAPKA NWRQALVISL DNRIQT...........  PEKLICPK T   MLQQNS   KK..... 548
N-GENE   MVS  STLRFSNQAV NMKRLRVFNM............GRSSTHYAIDYLPNN   CFVCTNYPW... S 588
L6       SI.PWGVKYEFKS CFLN SELRY HA............REAMLTGDFN L   N KW ELPFYK GEDDP 643
CON.

I2C-1    LPNDLFIKLKLLRFLDLSETSITKLPDSIFVLYNLETLLLSSCEYLEELPLQMEKLINLRHLDISNTRRL 669
I2C-2                      I R N KR      C             K.                  WH 660
RPM1       S..... N    A    EDS   S     CLVTMF  KY N  KTQ.VK    KNFH   V   ET NTKHSKIE 640
RPS2       I TGF MHMPV   V      F      EI L    KY VE YH SM GTK.ISV   QELGN RK K    LQR QF  617
N-GENE   F STF..E   M VH Q RH............... SLRH WTETKH PS .........   RI L WSK   632
L6       PLTNY..TM N IIVI EHSH   ADD......WGGWRHMMKMA R KVVR ASNYSLYG RVRL DCW F 705
CON.

I2C-1    KM..PLHLSRLKSLQVLVGAKFLVGGWRMEYLGEAHNLYGSLSILELENVVDRREAVKAKMREKNHVEQL 737
I2C-2     ..                   V     D     Q        VVK               P          728
RPM1     EL..  GMWK   K ......RY ITFR N GHDSNW Y............LGT ..  VP IWQ....... 682
RPS2     QTIPRDAICW SK E   .NLYYSTA   ELQSP   DEAEELGFAD   YLENLTTLGITVLSL.......... 677
N-GENE   T. RTPDFTCMPN EY..VN.. YQCSNL EVHHSLGCCSKVIG Y NDCKSLKRFPCVNVESLEYLGLR 696
L6       P. KSIEVLSMTAIEMDEVD..IGELKKLKT VLKFCPIQKI GGTFGMLKGL  L.........CL FN 762
CON.

I2C-1    SLEWSESISADNSQTERDILDELRPHKNIK.AVEITGYRGTNFPNWVADPLFVK.LVHLYLRNCKDCYSL 805
I2C-2                                  Q.E K I                 L .  K S         776
RPM1     ..........................L .DLQVMDC.......FN EDELI ..............N  702
RPS2     .............E LKTLFEFGAL   H Q.HLHVEECNELLYF LPSLTNHGRN RR SIKS H LEY  733
N-GENE    CDSL KLPEIYGRMKP..........E QIHMQGS I ELP.SSIFQYKTH TX LLWNM....KNLVA 751
L6       WGTNLREVV  IG LSSLKVLKTTGA EVEINEFPL LKELSTSSRIPNLSQLLD EV KVYD   GFDM 832
CON.

I2C-1    PALGQ.....LPCLEFLSIRGMHGIRVVTEEFYG............................RLS 837
I2C-2         .....    K   VK           ...........................         828
RPM1     GCMT .....  TRISLVMV RE   RDLC..................................... 725
RPS2     VTPADFENDW   S   V TLHSL NLTR...................................WG 763
N-GENE   LPSSICRLKS VS SVSGCSKLESLPEEIGDLDNLRVFDASDTLILRP...................P  801
L6        PASPSEDESSVWWKV...SKLKSLQLEKTRINVNVVDDASSGGHLPRYLLPTSLTYLKIYQCTEPTW P 899
CON.
```

FIG. 8B

```
I2C-1   SKKPFNSLVKLRFEIMPEWKQWHTLG.IGEFPTLEKLSIKNC......PELSLEIPIQFSSLKRLD.... 896
I2C-2       C E  E     T       A  .              I  ......                  FRVFGC 891
RPM1    ..DSL KIKRI  LSLTSIDEEEP E. DDL................................... 753
RPS2    NSVSQDC RNI CINISHCNKLKNVSWVQKL K  VIELFD ......R IEEL SEHE P......... 818
N-GENE   IIRL K II M RGFKDGVHFEFPPVAEGLHS  Y NLSY NLID  GG PE  .GSL    K  L..S 866
L6      GIENLEN TS EVN IFQTLGGDLD L.QGLRS  I R RKVNGLARIKG KDLLCSSTCK RKFYITEC 968
CON.

I2C-1   .........................ICDCKSVTSFPFSILPTTLKRIKISGCPKLKLEAPVGE..MFVEY 939
I2C-2   PVVF YDAQVLRSQLECMKQIEEIY R  N                 T D           C MS  L E 960
RPM1    ............................................. ATASIE  F AGXLERVPSWFNT 767
RPS2    ..................................VEDPT FPS  TLRTRDL E N........SILPSR 847
N-GENE  RNNFEHLPSSIA    QLGALGSLDLK  QRL QL ..E  PE NELHVDCHMA  FIHYLVTKRKKLHR 930
L6      POLIELLPCELGVQTVVVPSMAELT R  PRL.EVG..PMIRS PKFPM.....  KLDLAVANITKE D 1030
CON.

I2C-1   LSVIDCGCV..................DDISPEFLPTARQLSIENCHNVTRFLIPTATESLH.IRNC.. 987
I2C-2   F  EE     .....................          E R G      .          T  .   EN 1006
RPM1      ..QNLTYL..................GLRGSQLQEN .I   QTLPRLVWLSFYN YMGPR.L F... 821
RPS2    VKLD AHNDTMYNLFAYTMFQNISSMRR ISASDS SLTVFTGQPYPEKIPSWFHHQGWD SVSVN.... 996
N-GENE   DA GSLEELV  SLELELDDTSSGIERIVS SKLQKLTTLVV.....K PSLREIEGL E KSLQDLYL 1093
L6
CON.

I2C-1   .EKLSMACGGAAQLTSLNIWGCKKLKCLPELLPSLKELRLTYCPE........IEGELPFNLQILDIRYC 1048
I2C-2   V                D S                    Q  N   ........        K Y  D  1068
RPM1    ...........QGFQN K LEIVQM H T VV.....................  DGAM E  K YV A  860
RPS2    E  WWK LEKDQPNEE C........Y   RFV N................................... 909
N-GENE  ... PENWYIPDKFLGFAVCYSRS IDTTAH IPVCDDK...........MSRMTQK A....LSECDTE 1048
L6      EGCT LGRLPLEK KE D G   PD TE VQTVVAVPS   GLTIRDCPRLEVGPMIQS   KFPMLNELTLS 1163
CON.

I2C-1   KKLVNGRKEW.HLQRLTELWIKHDGSDEHIEHWELP....SSIQRLFIFNLKTLSSQHLKSLTSLQFLRI 1113
I2C-2                  K V Y      D         ....C  T  EV    I              Y C  1133
RPM1    RG EYVPRGIEN IN Q  HLI .V NQLV RIRGE....G VD ...........SRV HIPAIKHYFR 914
RPS2    ...................................................................... 909
N-GENE  SSNYSEWDIHFFFVPFAG   DTSKANGKTPNDYGIIRLSF GEEKMYGLR LYKEGPEVNA LQMRENSN 1118
L6      MVNITKED LEV GS E   DSLELTL DTCSSI RISF L KL  K TTLIVEVP LREIEG AE KS     1232
CON.

I2C-1   VGNLSQFQSQGQLSSFSHLTSLQTLQIWNFLNLQSLPESALPSSLSHLIISNCPNLQSLPLKGMPSSLST 1183
I2C-2   D     PI      I                  H      S          Q E FH         N       K 1203
RPM1    TD G FYV ...   ....................................................... 926
RPS2    ...................................................................... 909
N-GENE  EPTEHSTGIRR.TQYNNR    FYE ....ING......................................1144
L6      LYLEGCTSLERLWPDQQQ G   KN NVLDIQGCK  SVDH SALKTT  PPRARITWPDQ YR........1294
CON.

I2C-1   LSISKCPLLTPLLEFDKGEYWTEIAHIPTIQIDEECM 1220
I2C-2    L  G                 PQ        L  W YI 1240
RPM1    ..................................... 926
RPS2    ..................................... 909
N-GENE  ..................................... 1144
L6      ..................................... 1437
CON.
```

FIG. 8C

SUBSTITUTE SHEET (RULE 26)

```
558
L L P L D L Q F E Y . S K K L S K R V L H N I L P T L R S
L R A L S L S H Y Q . I E V L P N D L F I K L K L
L R F L D L S E T S . I T K L P D S I F V L Y N
L E T L L L S S C E Y L E E L P L Q M E K L I N
L R H L D I S N T R R L . K M P L H L S R
L K S L Q V L V G A K F . L V G G W R M E Y L G E A H N
L Y GSL S I L E L E N V . V D K R E A V K A K M R E K N H
V E Q L S L E W S E S I . S K D N S Q T E R D I L D E L R P H K N
I K A V E I T G Y R G T . N L P N W V A D P L F V K
L V H L Y L R N C K D C Y S L P A L G Q L P C
L E F L S I R G M H G I R V V T E E F Y G R L S S K K P F N S
L V K L R F E D M P E W K Q M H T L G I G E F P T
L E K L S I K N C P E L S LEL P I Q F S S
L K R L D I C D C K S V T S W P F S I L P T T
L K R H K I S G C P K L K LEL P V G E M F
V E Y L S V I D C G C V D D L P E F L P T
A R Q L S I E N C H N V T R F L I P T A
T E S L H I R N C E K L S M N C G G A A Q
L T S L N I W G C K K L K C L P E L L P S
L K E L R L T Y C P E I E GEL P F N
L Q I L D I R Y C K K L V N L R K E W H L Q R
L T E L W I K H D G S D E H L E H W E L P S S
I Q R L F I F N L K T L S S Q R L K S L T S
L Q F L R I V G H . . L S Q P Q S Q G Q L S S F S H L T S
L Q T L Q I W N F L N L Q S L P E S A L P S S
L S H L I I S N C P N L Q S L P L K G M P S S
L S T L S I S K C P L L T P L L E F D K G E Y W T E I A H
I P T H Q I D E E C M . .
                      1220
L . . L . I . . C . . L . . L P . . . . . . . .        27 repeats

L . . L . I . . C . . α . . α P . . . . . . . . . I2-family
L . . L . L S G C . . α . . L . . L . . . . . . . T-LR
L . . L . α . . . . . α . . α P . . α . . L . . RPS2
L . . L . L . . . . . L . . L P . . α . . L . . N-gene
L . . L D L S S N N L . G . I P S . L . . L . . Cf-9
```

Fig. 9

FIG.10

```
I2C-1   EEFYGRLSSKKPFNSLVKLRFEDMPEWKQWHTLGIGEFPTLEKLSIKNCPELSLEIPIQF  889
I2C-2                C E  E     T     A                I               880
I2C-3   ............................................................
I2C-4   Q            C E  E     TG    A         N            K        60
CON.    EEFYGRLSSKKPFN L KL FEDM  WKQWH LGIGEFPTLE LSI NCPELSL IPIQF

I2C-1   SSLKRLD............................ICDCKSVTSFPFSILPTTLKRIKIS  921
I2C-2        FRVFGCPVVFYDAQVLRSQLEGMKQIEEIY R  N            T D        940
I2C-3   ...........................................................N  1
I2C-4         QVRGCPVVFDDAQLFRSQLEAMKQIEALY R  N I          T E        120
CON.    SSLKR                              I DC S TSFPFSILPTTLK I I

I2C-1   GCPKLKLEAPVGE..MFVEYLSVIDCGCVDDISPEFLPTARQLSIENCHNVTRFLIPTAT  979
I2C-2                C MS  L EF  EE   ...       E R G    .             996
I2C-3   SARE           ..         N    E          K I TD Q    I          59
I2C-4           F      ..              L   A CK   MS   F              178
CON.        LK EAPV E  MFVE  SV  CGCVDDIS EFLP A  L I  C N TRFLIPTAT

I2C-1   ESLHIRNC...EKLSMACGGAAQLTSLNIWGCKKLKCLP....ELLPSLKELRLTYCPEI  1032
I2C-2    T      ENV                 D S          ....        Q  N     1052
I2C-3    T T E  ENV    V        M    I SE       ERMQ             SD     119
I2C-4    T T S  ENV    V        M L H LK        ERMQ         D Y SN     238
CON.    E L I NC   EKLS ACGGAAQ T L I  CKKLKCLP    ELLPSLKEL L CPEI

I2C-1   EGELPFNLQILDIRYCKKLVNGRKEWHLQRLTELWIKHDGSDEHIEHWELPSSIQRLFIF  1092
I2C-2            K Y  D                  K V Y      D        C  T  EV  1112
I2C-3            K Y S                       H      D           S T C  179
I2C-4          HK R SD                    V H      D        C   N RVY  298
CON.    EGELPFNL  L I   CKKLVNGRKEWHLQRLT L I HDGSDE IEHWELP SI  L

I2C-1   NLKTLSSQHLKSLTSLQFLRIVGNLSQFQSQGQLSSFSHLTSLQTLQIWNF.........  1143
I2C-2    I               Y C D    PI     I                 .........  1164
I2C-3    I               Y CFD    I                     R..........   228
I2C-4   MI               C S G    I... R                R  GNLQSLAES  355
CON.    N  TLSSQHLKSLTSLQ L   GNLS   QSQ LSSFSHLTSLQTLQI
                                                          1

I2C-1   ............................................................  1143
I2C-2   ............................................................  1163
I2C-3   ................NLQSLAALALPSSLSHLTILNFPNLQSLSESALPSSLSHLIIDDC  273
I2C-4   ALPSSLSHLTISRCPNLQSLAESALPSSLSHLNIYDCPNLQLLPESALPSSLSHLDISHC  415
CON.
                          2                    3

I2C-1   ........................LNLQSLPESALPSSLSHLIISNCPNLQSLPLKGMPSS  1180
I2C-2   ........................H     S       Q E FH        N         1200
I2C-3   PNLQSLSESALPSSLSHLDISNCP      S       S T YD       V          333
I2C-4   PNLQSLPESALLSSLSHLDISHCP      A         T  H    H  SE          475
CON.                            NLQSL ESALPSSLS L I  CPNL SL   GMPSS
                  4                         5

I2C-1   LSTLSISKCPLLTPLLEFDKGEYWTEIAHIPTIQIDEECM  1220
I2C-2    K L  G                PQ       L  W YI   1240
I2C-3    E A       K     G     PN     S Y  W R    373
I2C-4    K      S        N     N    S      WK     515
CON.    LS L IS C LL PLLEF KGEYW  IAHI  I ID
             6
```

FIG. 11

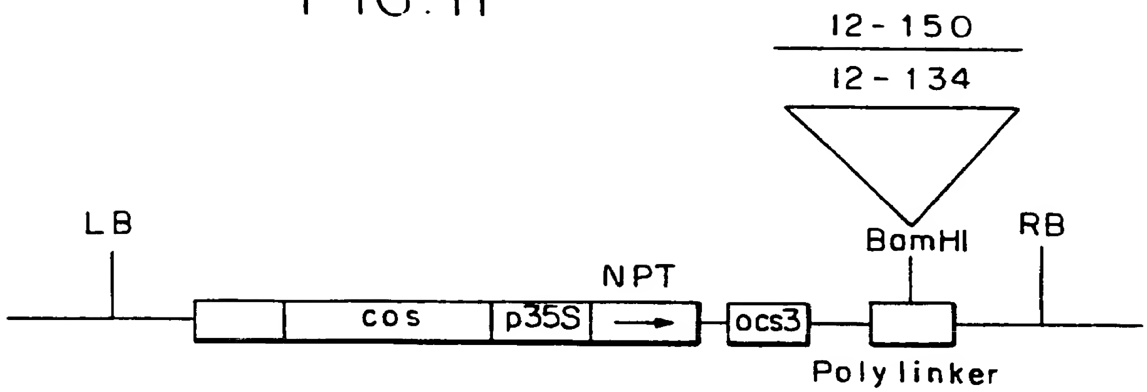


FIG. 12A
134A

FIG. 12B
134H

FIG. 12C

FIG. 12D
134B

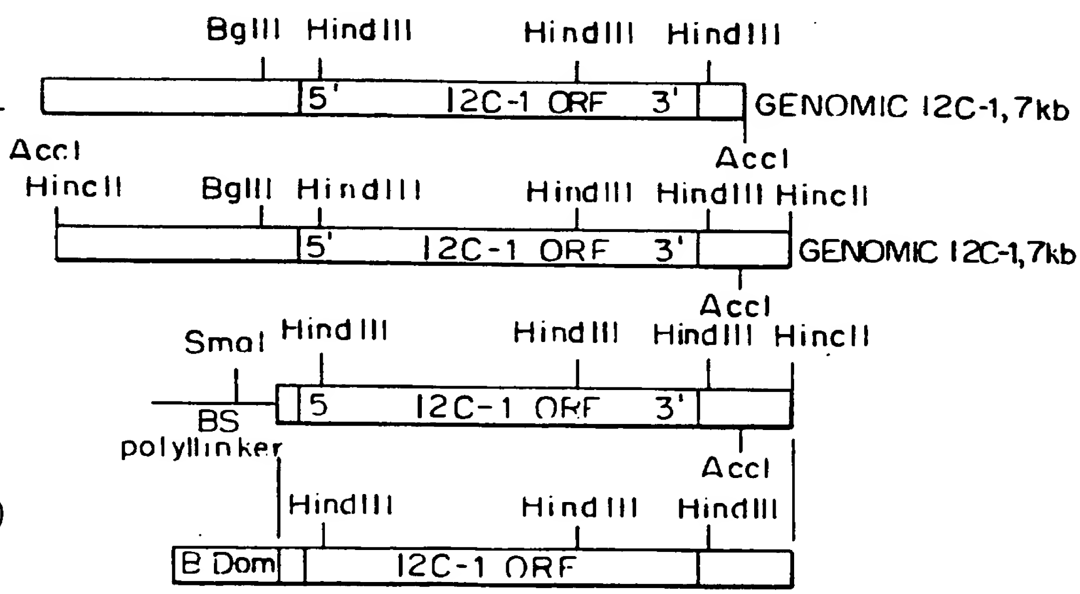
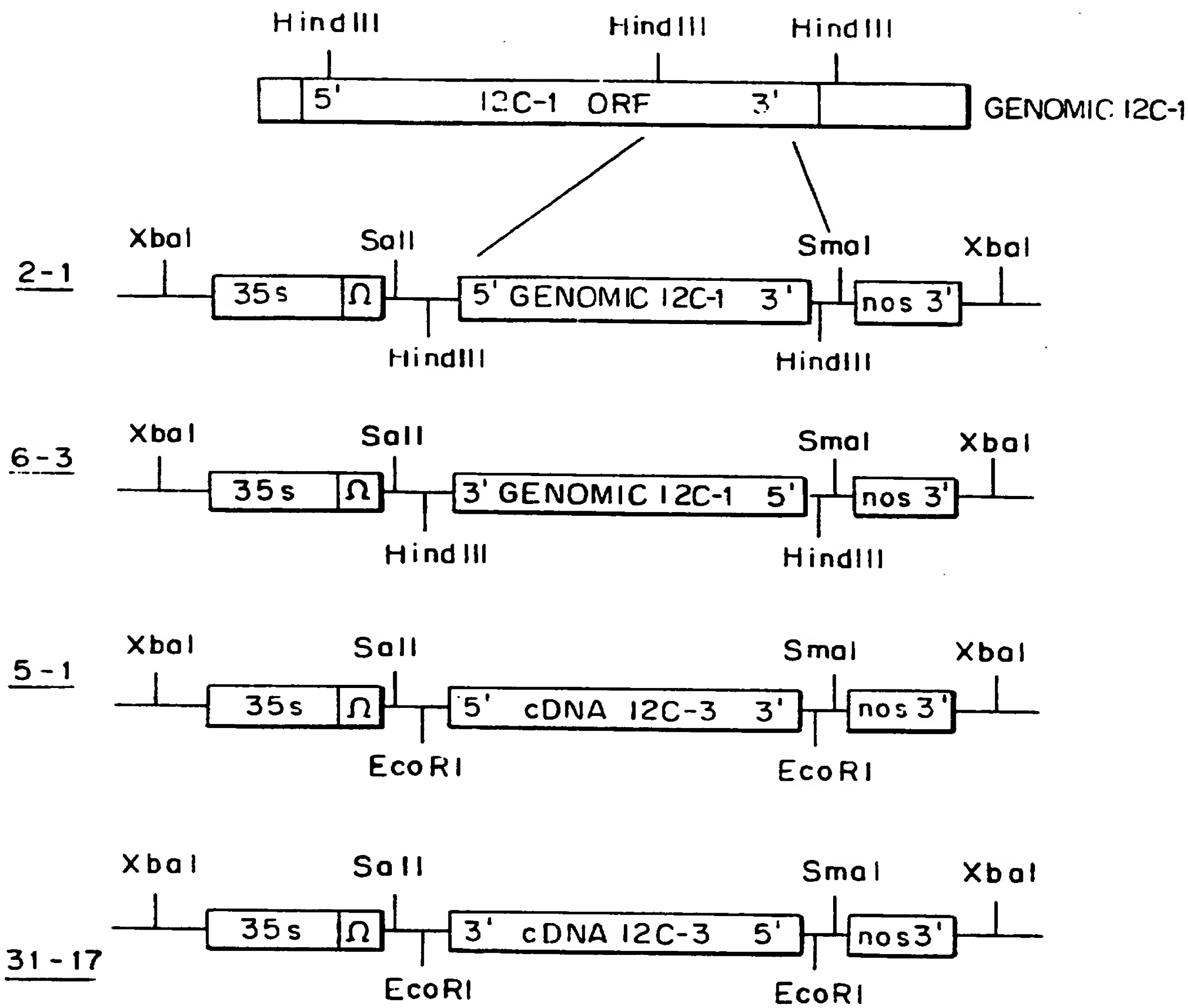


FIG. 13


HindIII
HindIII
HindIII
5'
12C-1 ORF
3'
GENOMIC 12C-1
2-1
XbaI
SalI
SmaI
XbaI
35s
Ω
5' GENOMIC 12C-1 3'
nos 3'
HindIII
HindIII
6-3
XbaI
SalI
SmaI
XbaI
35s
Ω
3' GENOMIC 12C-1 5'
nos 3'
HindIII
HindIII
5-1
XbaI
SalI
SmaI
XbaI
35s
Ω
5' cDNA 12C-3 3'
nos 3'
EcoRI
EcoRI
31-17
XbaI
SalI
SmaI
XbaI
35s
Ω
3' cDNA 12C-3 5'
nos3'
EcoRI
EcoRI

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/05272

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :A01H 1/04, 5/00; C12N 5/04, 15/29, 15/64, 15/82; C12Q 1/00;
US CL :800/205, DIG 44; 435/6, 240.4, 320.1; 47/58, DIG 1
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/205; 435/240.4, 320.1; 47/58, DIG 1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

SWISSPROT, MPSRCH
search terms: SEQ. ID NOS. 1-4

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --, E Y	US, A, 5,530,187 (LAMB et al.) 25 June 1996 (25/06/96), see entire document, especially columns 2-3.	1 -- 1-9
Y, P	US, A, 5,437,697 (SEBASTIAN ET AL.) 01 August 1995 (01/08/95), see entire document, especially columns 36-37.	1,10-13
X -- Y	SARFATTI et al. RFLP mapping of I1, a new locus in tomato conferring resistance against Fusarium oxysporum f. sp. lycopersici race 1. Theor. Appl. Genet. 1991, Vol. 82, pages 22-26, especially 25.	1 --------- 1,10-13

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 AUGUST 1996	03 SEP 1996

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Melissa Kimball
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

C (Continuati n). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BOURNIVAL et al. An isozyme marker for resistance to race 3 of Fusarium oxysporum f. sp. lycopersici in tomato. Theor. Appl. Genet. 1989, Vol. 78, pages 489-494, especially 493.	1,10-13
Y	BOURNIVAL et al. New sources of genetic resistance to race 3 of fusarium wilt of tomato. Plant Disease. 1991, Vol. 75, pages 281-284.	1,10-13
Y	US, A, 5,437,697 (SEBASTIAN ET AL.) 01 August 1995 (01/08/95), see entire document, especially columns 36-37.	1,10-13
Y, P	STASKAWICZ et al. Molecular genetics of plant disease resistance. Science. 05 May 1995, Vol. 268, pages 661-667, especially 666.	1,10-13